

**PERCEIVING NONHUMANS: HUMAN MORAL PSYCHOLOGY AND
ANIMAL ETHICS**

A Dissertation

by

TYLER JOSHUA KASPERBAUER

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Gary Varner
Co-Chair of Committee,	Clare Palmer
Committee Members,	José Bermúdez
	Linda Radzik
	Brandon Schmeichel
Head of Department,	Gary Varner

May 2014

Major Subject: Philosophy

Copyright 2014 Tyler Joshua Kasperbauer

ABSTRACT

There are currently very few discussions of moral psychology in the animal ethics literature. This dissertation aims to fill this void. My main contention is that many theories in animal ethics hold mistaken views about the moral psychology of human beings. These mistaken moral psychological views, I argue, limit these theories' ability to act as a guide in people's treatment of animals.

To develop my argument, I propose five criteria by which to assess the psychological plausibility of ethical theories, drawing from numerous recent developments in empirical moral psychology. I also draw a comparison between cases of physical impossibility in the "ought implies can" literature and cases of psychological difficulty, primarily as they arise in the literature on moral ideals. In both cases, I argue, limitations in individual resources constrain what ethical theories can ask of individuals.

I then investigate three different topics relevant to human moral psychology and normative evaluation of animals: attributing mental states to animals, the status of animals as disgust elicitors, and our empathic responses to animals. With respect to mental state attribution, I argue that the best research to date indicates that phenomenal mental states, like pain, determine our judgments of the moral considerability of animals. I also argue that the behavioral triggers we possess for attributing phenomenal states to animals are quite narrow—primarily animals that look and act like human beings. With respect to disgust, I examine research suggesting animals elicit disgust-based avoidance. I draw from research on dehumanization to argue that one way we cope with animals,

despite their disgust-evoking powers, is by attributing them mental states that evaluate them positively but simultaneously cement their status as inferior beings. In the case of empathy, I argue against the idea that empathy is psychologically central to expressing moral concern for animals. I examine six empirical claims made about empathy in the animal ethics literature and argue that all six are problematic to varying degrees.

I conclude by making suggestions for overcoming specific psychological obstacles identified throughout the dissertation. I also outline a research plan for constructing psychologically plausible theories in animal ethics.

TABLE OF CONTENTS

	Page
1. INTRODUCTION: MORAL PSYCHOLOGY IN ANIMAL ETHICS.....	1
1.1 Using Psychology to Inform Ethics.....	3
1.2 The Goals of Ethics.....	5
1.3 Moral Psychology in Animal Ethics.....	9
1.4 Section Outline.....	11
2. PSYCHOLOGICAL PLAUSIBILITY IN ETHICS.....	13
2.1 Basic Psychological Constraints.....	15
2.2 Moderate Psychological Constraints.....	36
2.3 Conclusion: Implications of Psychologically Difficult Ideals.....	49
3. MENTALIZING ANIMALS: IMPLICATIONS FOR ANIMAL ETHICS.....	53
3.1 Moral Considerability and Moral Significance.....	55
3.2 The Phenomenal Account.....	66
3.3 Phenomenal States and Moral Significance.....	81
3.4 Ethical Implications of the Phenomenal Account.....	84
4. ANIMALS AND DISGUST.....	92
4.1 Animals as Disgust Elicitors.....	93
4.2 Animal Reminder, Mortality Salience, and Terror Management.....	104
4.3 Dehumanization and Infrahumanization.....	112
4.4 Ethical Implications of Disgust Toward Animals.....	116
5. REJECTING EMPATHY FOR ANIMAL ETHICS.....	127
5.1 What is Empathy?.....	129
5.2 Empirical Claims of the Centrality Thesis.....	139
5.3 Ethical Implications of Empathy.....	164
6. CONCLUSION: TOWARDS PSYCHOLOGICAL PLAUSIBILITY.....	169
6.1 Summary of Main Claims.....	169
6.2 Managing Psychological Limitations.....	172
6.3 Conclusion: Towards Psychological Plausibility in Animal Ethics.....	181
REFERENCES.....	183

1. INTRODUCTION: MORAL PSYCHOLOGY IN ANIMAL ETHICS

Here is a simplified snapshot of the current state of moral psychology in philosophy: First, some philosophers focus primarily on empirical studies within psychology, neuroscience, and other sciences. They attempt to achieve clarity on the theoretical issues at stake, adjudicate competing claims, advance hypotheses, and synthesize and interpret empirical research for philosophical audiences. These are all typical tasks for philosophers, firmly established in the history of the discipline, but they are also generally recognized as tasks for psychologists—or descriptive moral psychology. Secondly, other philosophers, in the process of developing theories of the right or the good, advance particular theories of moral judgment, intentionality, motivation, the relationship between beliefs and desires, and other aspects of moral psychology that are not as easily studied in the lab. Thirdly, an increasingly popular approach is to use the first category to inform or critique the second. There is no particular name for this third type of approach, as it straddles both empirical and theoretical aspects of moral psychology.

This dissertation adopts the third approach. However, it also focuses on animal ethics, which tends to emphasize traditional normative and meta-ethics without regard for moral psychology. Recent advances in empirical moral psychology are mostly absent from prominent discussions of, for example, the moral status of nonhumans, intensive animal agriculture, species rights, and a variety of other issues. I aim to change this trend. My explicit aim in this dissertation is to use studies from the cognitive sciences, primarily psychology, to improve and inform the development of theories in animal

ethics. I seek to identify and construct ethical theories about animals that are morally defensible *and* psychologically plausible.

My main contention in this dissertation is that many theories in animal ethics hold mistaken views about the moral psychology of human beings. These mistaken moral psychological views, I argue, limit these theories' ability to act as a guide in people's treatment of animals. To develop my argument, I propose five criteria by which to assess the psychological plausibility of ethical theories, drawing from numerous recent developments in empirical moral psychology. I then investigate three different topics relevant to moral psychology and normative evaluation of animals: attributing mental states to animals, the status of animals as disgust elicitors, and our empathic responses to animals. In the concluding section, I make suggestions for overcoming psychological obstacles pertaining specifically to the perception of animal bodies. I also outline a research plan for ethicists and moral psychologists interested in our treatment of nonhumans.

This section will provide a basic summary of the role of moral psychology in ethical theory generally and animal ethics specifically. Each of the following sections will take on this task in their own way, but here I will try to capture general moral psychological themes so as to help frame the rest of the dissertation. This is intended to orient my dissertation within the moral psychology and animal ethics literature. I will discuss three general issues: how psychology can be argued to impact ethical theory, how to think about the goals of ethical inquiry, and how animal ethicists have approached moral psychology.

1.1 Using Psychology to Inform Ethics

Many ethicists seem to think of moral psychology as ancillary to moral theory. The claims I advance can be understood as reversing this relationship. Consider two potential projects someone might adopt in using psychology to inform ethics.

The first I'll refer to as *descriptive accuracy*. Suppose that most normative projects in contemporary ethics are working with an inaccurate picture of human beings, whatever that might mean. Ethicists could see this as bad in its own right, regardless of how it might impact their normative projects. By analogy, denying what human beings are really like could be considered similar to denying the facts of history, mathematics, or any number of other disciplines. Even if getting clarity on facts about human beings *wouldn't affect our normative projects at all*, we could still see descriptive accuracy across disciplines as a worthwhile goal.

A second possible project I'll call *deception avoidance*. Someone could argue that ethicists are frequently too biased towards their own projects, and construe human psychology however they wish for normative ends. Kantians say we are rule-following creatures, Humeans say we are fundamentally driven by emotions and desires, virtue theorists emphasize character traits, and so on. Here, empirical moral psychology constitutes the loyal opposition to ethical theory. Moral psychology keeps ethicists in check by making sure they are staying within their descriptive bounds when making normative (or meta-ethical) claims. This doesn't inform ethical projects so much as it keeps our biases in check.

My project is not to merely “fact-check” normative theories nor help keep their biases under control. Rather, the psychological limitations I identify are also aimed at altering the content of ethical prescriptions. My aim is to show why mistaken descriptive views about human moral psychology are undesirable, and why uncovering them should lead us to amend our ethical theories. To be clear, I do not wish to dismiss moral philosophy, or replace moral philosophy with social science, but rather to explain how moral philosophy can operate within the bounds of human psychology.

Most ethicists seem to think that being mistaken on descriptive issues is only trivially bad. If I have the *right* moral theory, then what relevance is moral psychology? If I have *correctly* identified a moral duty, why does it matter what I say about human beings’ capacity to realize that duty? For instance, suppose I claim that our highest moral duty is to increase the amount of pleasure and decrease the amount of pain in the world. And suppose I derive from this that our current treatment of nonhuman animals must change dramatically. Assuming these claims are right, and the reasoning sound, what relevance is whatever moral psychology that follows? For instance, why should it matter if people are unmoved by the pain felt by animals?

The main problem, which I address in detail in section 2, is that human agents are not always capable of responding in the right way to moral theories—however correct they might seem to philosophers and professional ethicists. This inability to properly respond is varied: some will be confused by what ethical theories are proposing, some will reject the proposals as illegitimate (for whatever reason), and some won’t be motivated by ethical theories. I do not wish to take any specific stance here on what

causes the improper response (though I will in later sections). What is clear is that when ethical prescriptions fail to have an impact, they also fail to serve as guides. The prescriptions offered by ethical theories are not “live options” for many people. They might offer paths to action—and so are at least nominally what ethicists would describe as “action guiding”—but these paths might not be available to large segments of society.

Perhaps many ethicists are willing to accept this consequence. If the goal of ethics is to identify moral truths, for instance, then the only people whose responses count are those capable of perceiving moral truths. That ethical prescriptions are not “live options” for many people might be considered irrelevant to the project of ethics. This can be understood by analogy to science. We might say the project of science is to identify truths about the world, regardless of what people believe. In science, we would never suggest that the truth of things is affected by how people, in general, respond to scientific proposals. Why think ethics is any different?

1.2 The Goals of Ethics

I think the projects of ethics and science are similar in many ways. One way in which they are not the same, however, is in addressing human psychology. Whereas reasons for accepting scientific claims need not act as a guide for normal human beings, reasons for accepting ethical claims should.

The idea that ethical theories should be guiding in some way has been endorsed by a wide variety of different philosophers. Some prominent moral and political philosophers, for instance, have criticized ethical theories for ignoring the action options

human agents actually face, and instead focusing on the action options agents would have in ideal conditions (e.g., Gheaus, 2013; Velleman, 2013). That ethical theories should be guiding in some way is a particularly widespread view among applied ethicists. For instance, among environmental ethicists, Ronald Sandler (2003, 2004, 2007) argues that action guidance is a necessary component of any *adequate* environmental ethic. On Sandler's view, an adequate environmental ethic "must recommend a course of action...in concrete situations regarding individual or communal interactions or relationships with the natural environment" (2004, p. 479). An adequate environmental ethic, he elaborates, "must be efficacious in promoting solutions to real world environmental problems. It must help bring about, not merely justify, environmentally sustainable practices, policies, and lifestyles" (2007, pp. 107-108). Though I am not arguing for action guidance specifically, this sort of view illustrates the interest many ethicists have in producing theories that will have concrete effects on ordinary human beings.

It is furthermore a common view among ethicists that ethical inquiry is *distinctly* unlike scientific inquiry in that it should act as a guide for normal human beings. However, I disagree with the arguments often offered for this. A classic expression of this idea can be found in Korsgaard (1996, lecture 1). Korsgaard argues that normative ethics is, by its nature, first-person oriented: What ought *I* do? On the first-person perspective, normative ethics, unlike science, aims to provide individuals with reasons for action. This is like science in one respect, however, in that ethicists' answers to this first-person question are taken to be independent of general human psychology.

I think normative ethics should ask a slightly different, third-person, question: What ought *people* do, in general? The main consideration in support of this is that answers at the first-person level might not provide guidance for anyone else. This is a basic lesson ethicists could learn from cognitive science. Cognitive scientists generally seek out statistical regularities among groups of individuals. Rather than ask what individuals should do, ethicists should ask what various psychologically similar groups of human beings should do. This is the best way to ensure that ethical prescriptions generalize beyond one's own perspective. If normative ethics is first-person oriented, then what use are ethicists, whose ethical views might apply only to their particular perspective?

Third-person ethics focuses instead on providing recommendations that take into account what people are capable of, and what actions their psychological profiles¹ are best suited for, while still providing substantive normative content. Third-person ethics is unlike scientific projects in that it takes people's beliefs to influence the content of the questions at hand (i.e., ethical prescriptions). However, it is broadly scientific in that the object of study—human psychology—is understood in terms of statistical regularities. Given sufficient empirical data, ethical prescriptions can be tailored according to prominent psychological features of the general population. This may ultimately culminate in an individual (or individuals) asking the first-person question “What ought I do?” but the general framework for ethical decision-making, as promoted by ethicists, is determined in accordance with broader patterns existing among groups of human

¹ I use “psychological profile” as shorthand for all of a person's psychological dispositions.

beings. This is not to say that these broader psychological patterns *settle* any ethical issues, but that ethicists must understand these patterns in order for their theories to have “psychological grip” on a variety of psychological profiles.

To illustrate the point I am trying to make about third-person ethics, consider two different types of potential normative projects. Suppose that the *correct* metaphysical account of responsibility holds that none of us are responsible for our actions. On one sort of normative project, ethicists would recommend that those in charge of enacting laws regulating punishment and blame should take this into account. On another sort of normative project, ethicists would assume that people are likely to reject the idea that we ultimately are not responsible for our actions, however strong the metaphysical case for this view. Here, those in charge of enacting relevant laws concerning punishment and blame would need to take into account both the metaphysics and the psychology of responsibility.

While first-person ethics might be adequate for the first normative project, it seems hopeless for the second. Ethicists would not be able to assume that answers from first-person theorizing would provide guidance. What any particular ethicist, politician, or lawmaker would do in dictating punishment and blame—from the first-person perspective—would potentially be irrelevant to anyone else. In this case, and arguably in many others, first-person ethics appears to serve as a guide only by assuming significant uniformity in psychological profiles.

1.3 Moral Psychology in Animal Ethics

For the most part, animal ethicists have not engaged with any of the central debates in contemporary moral psychology. However, one relevant debate within the field is over the role of reason in changing our attitudes to and treatment of animals. Like normative ethicists generally, animal ethicists too have adopted a broadly first-person approach. Perhaps the most prominent moral psychological claim found in the animal ethics literature, commonly attributed to Peter Singer and Tom Regan, is that human agents need only use reason to recognize their moral duties to animals. From the first-person perspective, that is, individuals should be able to reason their way through any psychological obstacles to see what is required of them.

Consider Tom Regan's (1991) argument against care ethics. A key feature of care ethics is that our moral duties arise out of our close relationships with others. We owe more to our offspring, for example, than we do to strangers. Regan (1991, pp. 96-98) claims that this is not only morally objectionable but also psychologically implausible. He argues that human beings, as rational agents, are psychologically constrained by morally arbitrary principles—we recognize and want to avoid them. And showing partiality towards others, he argues, is one of these morally arbitrary principles. For example, in the case of pain, Regan does not see how we could say that the pain of one person matters more than the pain of another (including other animals). Pain is undesirable always and everywhere, and rational agents should realize this. If someone's emotional attachments are causing unreasonable partiality, Regan suggests that we work to bring our sentiments in line with our reason—our reason can tell us who to care for.

Regan's views on the function of reason in producing moral change are quite common in the animal ethics literature. That is, many animal ethicists assume that reason can alter behavior and reign in our emotions to the degree required by their theories. Though I think these claims are misguided, I do not intend to investigate them in any detail. What is significant, and somewhat surprising, is how rarely the proponents of these claims provide evidence in their support. Hundreds of studies, some of them including classic, widely cited experiments, could be cited to illuminate the role of reason in producing (or failing to produce) moral change, but they are entirely absent from the major publications in the field (the single exception I'm aware of is Varner, 2012). Animal ethicists seem to adopt a position similar to that summarized above: moral psychology does not matter, so long as we correctly identify what we owe to animals.

On such an approach, Regan's assertions seem puzzling in multiple ways. For instance, he is making falsifiable moral psychological claims (e.g., reason resists moral arbitrariness, partiality can be reduced through reason, pain is perceived to have equal and fundamental moral importance wherever it is found), and is also arguing against another normative theory's moral psychology (the partiality presupposed by care ethics is easily modifiable). However, the main motivation behind his approach seems to be the objectionable consequences that follow from the moral psychology of care ethics. Suppose he is right about this, and care ethics does indeed present an undesirable picture of human psychology. The burden would still be on Regan, it would seem, to demonstrate that his own moral psychological claims are at least minimally plausible.

Without this, all we are left with is his first-person drive to be more rational and bring sentiments in line with reason. It is not clear, however, why we would think this applies to human beings more broadly.

1.4 Section Outline

The moral psychology of other theories in animal ethics will be discussed throughout other sections of this dissertation. Before investigating claims made by animal ethicists specifically, I must say more about why the psychological plausibility of ethical theories has implications for ethical theories. Section 2 takes on this task, arguing for both minimal and moderate criteria for assessing the psychological plausibility of ethical theories. This will fill in and justify some of the claims made in this section.

Sections 3, 4, and 5 look at specific moral psychological issues with implications for the psychological plausibility of theories in animal ethics. Section 3 focuses on mental state attributions made to animals, or *mentalizing*. Though animal ethicists frequently use mental states to underpin animals' moral status, rarely is the relevant psychology discussed. I present arguments against the use of certain mental states for judgments of both moral considerability and moral significance.

Section 4 focuses on the role of disgust in producing aversion to animals. Ethicists rarely consider what obstacles exist to improved treatment for animals. Disgust, I argue, is one of the most prominent obstacles in this regard. I further argue for a specific way of understanding animals' prominence in human lives, despite their disgust-evoking powers.

Section 5 argues against the importance of empathy in ethical theorizing related to animals. Empathy has been central to a variety of ethical theories, both in the context of animals and in normative ethics generally. I argue against six empirical claims made about empathy and its role in producing moral concern for animals, concluding that empathy should not be ascribed a privileged role in our moral theorizing about animals. If the goal is to increase moral concern for animals (which includes improving our treatment of them), ethicists should look to other emotions besides empathy.

In sections 3, 4, and 5 I evaluate the pertinent theories according to their psychological plausibility. I also suggest ways of making theories in animal ethics more psychologically plausible. In section 6, I summarize my conclusions and outline the best strategies for animal ethicists in going forward. Many theories in animal ethics can be improved, rather than discarded entirely, by paying greater attention to what empirical research tell us about human psychology.

2. PSYCHOLOGICAL PLAUSIBILITY IN ETHICS

In the last decade, moral psychologists in both philosophy and psychology have made a number of sustained attacks against key assumptions and positions in philosophical ethics. Their basic claim is that most work in ethical theory is psychologically implausible because human psychology is radically different from what these theories assume or require. These criticisms have been largely empirical, consisting either of experimental explorations into morality or drawing heavily from such research. Despite these dedicated efforts, and much supporting data, many professional ethicists remain unconvinced that their theories are deficient or in need of revision. Indeed, in animal ethics, even citations of these critiques are absent from the major professional journals.

In this section, I aim to clarify the general methodology employed by these critiques and demonstrate that the methodology has merit. Ethicists might disagree with specific empirical critiques of their theories, but, I argue, they should acknowledge the general implications of moral psychology for ethics. The focus of my inquiry will be what I refer to as the *psychological plausibility* of ethical theories. The concept of psychological plausibility will be further refined throughout the section, as I apply it to particular empirical critiques. However, a rough operational definition of this concept is as follows:

An ethical theory is psychologically plausible if human agents are capable of responding to it in the right ways.²

² This definition is admittedly vague. This is necessary, however, in order to capture what I take to be the core theme across a wide variety of recent empirical critiques of ethical theories. For

This idea, so defined, is central to a wide range of empirical critiques of ethical theories.³

In further clarifying the concept of psychological plausibility, my aim is to assist ethicists in both evaluating these sorts of critiques and ensuring that their theories are psychologically plausible.

The first part of this section will discuss Owen Flanagan's Principle of Minimal Psychological Realism, one of the few attempts in the literature at providing basic psychological constraints on ethical theories. I will argue that there are certain deficiencies in Flanagan's Principle, and then lay out five criteria of minimal psychological plausibility to replace his. These criteria, I argue, provide independently necessary and jointly sufficient conditions for psychologically plausible ethical theories. If an ethical theory fails to meet any of these criteria, then it must be considered psychologically implausible.

There are also important issues raised by what I will call moderate (as opposed to minimal) psychological plausibility, or theories that meet my minimal conditions but are psychologically difficult to achieve. The second half of this section will explore the normative implications of psychological constraints at this moderate level of

instance, one might expect that this pertains specifically to "ordinary" human agents, or some similar qualifier specifying which human beings are in question. However, as I discuss below, some ethical prescriptions (e.g., those from virtue ethics) are explicitly aimed at excellent human agents, and so I have left the type of human agent unspecified. Similarly, different theories have different expectations regarding the "right ways of responding" (e.g., in action or in attitude), and so that too has been left unspecified. Later refinements of the concept of psychological plausibility will make this clear.

³ A slightly more refined conception of psychological plausibility is in terms of the flexibility of human psychology, and whether it can change when ethical theories dictate that it must. Though it is rarely stated explicitly, many empirical critiques of ethical theories suspect that human psychology is not as it needs to be to fit certain ethical theories' goals nor can it be modified to meet those goals at some time in the future.

psychological plausibility. Traditionally, ethicists have been dismissive of moderate psychological constraints, since, it is often argued, these constraints can be overcome by incrementally altering relevant social and political conditions. I reject this idea and argue that moderate psychological constraints do have normative implications for ethics. I draw a comparison between cases of physical impossibility in the “ought implies can” literature and cases of psychological difficulty, primarily as they arise in the literature on moral ideals. In both cases, I argue, limitations in individual resources limit what ethical theories can ask of individuals. I conclude by considering some objections and providing an example from animal ethics to illustrate broader implications of my argument.

2.1 Basic Psychological Constraints

2.1.1 Flanagan’s Principle of Minimal Psychological Realism

The classic account of psychological plausibility in morality is Owen Flanagan’s (1991) *Varieties of Moral Personality*. Here, Flanagan introduces his Principle of Minimal Psychological Realism (PMPR). This principle states, “Make sure when constructing a moral theory or projecting a moral ideal that the character, decision processing, and behavior prescribed are possible, or are perceived to be possible, for creatures like us” (1991, p. 32). Many who have written about the relationship between empirical moral psychology and ethics have cited the PMPR favorably (e.g., Appiah,

2008; Casebeer, 2003; DesAutels, 1998; Goldman, 1993; Johnson, 1996). None of the advocates of the PMPR have discussed its implications in any detail, however.⁴

One implication Flanagan sees for the PMPR is that ethicists must provide an account of the moral psychology required (or entailed) by their theories. The challenge for ethicists posed by the PMPR, according to Flanagan, is that “every moral conception owes us at least a partial specification of the personality and motivational structure it expects of morally mature individuals” (1991, p. 35). That is, every theorist owes an account of what moral psychological profiles are demanded by his or her theory. Flanagan thinks that specifying these demands should take into consideration what is psychologically possible for human beings.

Flanagan thinks his formulation of the PMPR is both descriptive and prescriptive. That is, he aims to identify a common feature of ethical theories (they in fact often aim, even if implicitly, to adhere to the principles contained in the PMPR) as well as to provide guidelines for proper use of ethical theories (that they *should* adhere to the principles contained in the PMPR, if they do not already). Two examples he cites of theories that violate the PMPR are specific understandings of virtue theory and act utilitarianism. Virtue theory violates the PMPR if it requires moral agents to hold every relevant virtue, since some virtues are inconsistent with each other (e.g., vivaciousness

⁴ Sabini and Silver (2005), in discussing virtue ethics, note that philosophy has been less than explicit in its adherence to something like the PMPR. However, they express the implications only in general terms: “Ethics is otiose if it prescribes behavior that people cannot perform or, more generally, if it urges people to be the sorts of people they cannot become. But if empirical research discovers that no one does (or is) what virtue ethics say they should do (or be), then that is, at least, prima facie evidence that people cannot do (or be) what virtue ethics requires” (pp. 537-538).

and serenity), and thus even the most virtuous agent can only possess a certain set of the virtues. With respect to act utilitarianism, a version of the theory that violates PMPR is one that requires moral agents to maximize outcomes in every “action opportunity.” We could conceivably seek out the best actions every second of our life, but such a demand would seem to be physically, or computationally, impossible. Since it is impossible for human beings to achieve the requirements of both of these theories, they violate the PMPR, and must be discarded.

Ultimately, however, Flanagan thinks most ethical theories satisfy the PMPR. The examples he cites pertaining to virtue ethics and act utilitarianism are purposely exaggerated for illustration. Moreover, Flanagan thinks the PMPR serves as only a very basic constraint, and is generally unsatisfactory as a *guide* for moral philosophy, for two main reasons: 1) there are many unactualized but possible psychological profiles that we disapprove of, and so wouldn’t want to see actualized, and 2) there are numerous unactualized but possible psychological profiles that we approve of and *would* like to see actualized. The main project of ethics, Flanagan seems to suggest, is to eradicate the bad profiles and seek out the good; on this, the PMPR offers no assistance. The role of ethics, one might argue, is to expand our conceptions of what is possible beyond our current notions. Flanagan approves of this goal and so applies the PMPR only as the most basic constraint on ethical theorizing.

While I agree with many of Flanagan’s conclusions, his account of the PMPR is too lenient, and concedes too much to ethical theories that should be considered psychologically implausible. It is ineffective even as a basic constraint on ethical

theorizing. The PMPR prompts ethicists to consider the psychological entailments of their theories, and whether they are in fact realizable, but this is not sufficient for assessing minimal psychological plausibility.

For instance, the definition of PMPR does not say anything about how we should assess what is possible for human beings, or how overall possibility relates to what is possible for currently existing humans. In fact, later in the book, Flanagan broadens his conception of possibility considerably:

PMPR does *not* demand that the character and motivational structure required by an acceptable ethical theory must now be realized, or have once been realized, or be realized on average in actual persons. PMPR requires only that the recommended ideals be possible under some conceivable social arrangement or other. (p. 201)

Under this augmentation, a theory need only specify “some conceivable social arrangement” to become psychologically plausible. This illustrates the PMPR’s inability to function as a constraint on ethical theorizing. It would seem to rule out only those theories that are, strictly speaking, totally impossible (such as act utilitarianism under Flanagan’s construal). This makes the normative aspect of the PMPR—prompting ethicists to make sure their theories are realizable—mostly irrelevant. An ethicist need only specify some conceivable social arrangement and thus satisfy the PMPR.

2.1.2 Criteria of Psychological Plausibility

I have suggested that Flanagan’s PMPR is too weak and barely filters out any kind of ethical theory as psychologically implausible. Now I will present and defend five alternative, stronger, criteria of psychological plausibility to replace Flanagan’s PMPR.

These criteria provide independently necessary and jointly sufficient conditions for making a theory psychologically plausible. If a theory fails to meet any one of these, I will argue that it must be judged psychologically implausible. I will also briefly review some of the most prominent recent empirical criticisms of ethical theories in order to support the criteria I propose.⁵

The first criterion I will refer to as *cohesion*, or whether one's theory is supported by what we currently know about the psychology of human beings. This criterion maintains that the psychological assumptions of ethical theories should be assessed according to scientific understandings of human psychology. This is perhaps the most general strategy one can find in recent criticisms of ethical theories, and it is also arguably the most common. Many critics think ethicists should pay more attention to what human beings are really like, and make their theories cohere with what cognitive scientists have discovered about the processes undergirding morality. When ethical theories are subjected to empirical scrutiny, ethicists should respond by demonstrating how the best empirical evidence supports their theories, rather than denying that such evidence is relevant. This can be stated as:

- 1) *Cohesion*: An ethical theory is psychologically plausible if it is supported by current scientific understandings of the psychology of human beings.

Of course, some ethicists are likely to reject this criterion. Why, they may ask, should scientific research impact ethical theories? What can such information teach ethicists?

⁵ Clipsham (forthcoming) argues against some of the critiques I will be discussing on the grounds that they do not settle any substantive issues in ethical theory. I am only outlining basic constraints on ethical theory, not trying to settle any issues, and so Clipsham's claims will not be addressed further.

I will illustrate the importance of this criterion by discussing two experiments that have prompted ethicists to look more closely at the cognitive sciences. The first experiment comes from Joshua Greene, who has conducted a variety of studies using fMRI to determine which parts of the brain are responsible for different aspects of moral judgments. In Greene's most widely known experiment, different types of moral reasoning were compared in the context of various moral dilemmas, such as the classic trolley problem (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001). In one trolley case, participants were asked whether it was permissible to pull a lever to divert a trolley away from five people on one track, toward another track on which there is only one person. In another trolley case, participants were asked whether it was permissible to push a fat man off a bridge in order to stop the trolley before it reached the five people on the track. Participants' brain activity indicated that there was greater activation in emotional areas of the brain in the fat man case than in the lever case. There was also greater activation in working memory areas of the brain in the lever case than in the fat man case. Greene et al. argue that these results provide evidence of two dissociable processes in our moral psychologies, one that is quick, automatic, and emotional, and another that is slow and rational.

There are of course many criticisms one could make of this experiment. However, we can put these aside to get at the issues relevant to psychological plausibility. Greene argues that his results show that the element of *personal force* biases our moral judgments. In Greene et al. (2001), as well as subsequent studies (Cushman & Greene, 2012; Greene, 2008; Greene et al., 2009), people respond more emotionally

when harms are caused by direct force, often by physically pushing another person, than when harms are caused more indirectly (e.g., by switching the trolley lever). Since the amount of personal force is irrelevant to the resulting harm (whether one or five people are harmed), one way of reading this is that our emotionally based judgments are attuned to factors that are morally irrelevant.

One way in which this research can be used to undercut psychological assumptions made by ethicists is in defense of consequentialism. The fat man case is often cited as a counterexample to consequentialist theories, which generally maintain that the right choice in both trolley cases is to save the five, regardless of the steps required to do this (Sinnott-Armstrong, 2006). If Greene is right, and personal force introduces biases into our moral judgments, then the fat man case should not be relied upon in arguing against consequentialism.⁶

However, there has been much debate over whether personal force is indeed morally irrelevant as well as more generally whether our emotional responses to personal force are epistemically reliable (Berker, 2009; Cushman & Young, 2009; Kamm, 2009; Kumar & Campbell, 2012; Sauer, 2012a). It is clear that the element of personal force influences our evaluation of outcomes, and it does so outside of cognitive control, but perhaps this is not problematic. I do not intend to enter into this debate. We can see that Greene's research poses a challenge to ethicists, regardless of whether his

⁶ As he explains the importance of moral psychology, "science does matter for ethics, not because one can derive moral truths from scientific truths, but because scientific information can challenge factual assumptions on which moral thinking implicitly depends" (Greene, 2008, p. 67).

overall argument is ultimately successful. Those who would use the fat man case to argue against consequentialism, for instance, must explain why those intuitions are epistemically reliable (i.e., whether they identify the right moral aspects of the trolley problem), given that there is now considerable evidence to suggest that they are not. Providing a satisfactory response to this challenge would require ethicists to more closely investigate the relevant empirical research, as suggested by the cohesion criterion.⁷

Another well-known study, illustrating a similar phenomenon, comes from Amos Tversky and Daniel Kahneman (1981), who asked participants to consider a hypothetical case in which 600 people are at risk from a new strain of some disease. In one condition, participants were asked to make a decision that would yield results framed in terms of “saving,” as follows:

- A. 200 people would be saved.
- B. There is a 1/3 probability that 600 people would be saved, and 2/3 probability that no one would be saved.

In another condition, participants received choices framed in terms of “dying”:

- C. 400 people would die.
- D. There is a 1/3 probability that no one would die, and a 2/3 probability that 600 people would die.

The outcomes of each condition are in fact identical. A and C lead to the same outcomes, as do B and D. They are just framed differently. Given that the outcomes are the same, one might expect people to have consistent preferences across the two cases. However,

⁷ To be clear, looking to the relevant empirical research is necessary but not sufficient for addressing the epistemic reliability of our moral intuitions; it will not *settle* the issue. However, providing sufficient answers to moral questions is not required by the cohesion criterion.

participants in the first condition overwhelmingly preferred A over B, while participants in the second condition overwhelmingly chose D over C, indicating that their preferences were not consistent. The generally accepted explanation for these results is that the language of “saving” biases people *toward* outcome A, and the language of “dying” biases people *against* outcome C.

There are clear parallels here to Greene et al.’s study. In both experiments, seemingly irrational processes influence people’s evaluation of outcomes. In Greene et al., the irrational processes seem to violate the normative standards of consequentialism (though I don’t wish to stake on a claim on whether they in fact do). Tversky and Kahneman (1986) argue that people’s responses to the disease case violate what they call the *principle of invariance*. This principle states, “different representations of the same choice problem should yield the same preference. That is, the preference between options should be independent of their description” (p. 253). This principle is aimed at the normative guidance provided by decision theory. The principle of invariance should be normatively guiding, they argue, because “Two characterizations that the decision-maker, on reflection, would view as alternative descriptions of the same problem should lead to the same choice—even without the benefit of such reflection” (p. 253). That is, people should be able to avoid framing effects, especially ones as simple as those exploited in the disease case, and evaluate a decision based solely on its outcomes. Ultimately, however, Tversky and Kahneman argue that decision theory cannot provide normative guidance because the evidence suggests people are incapable of avoiding these effects, even when they have time to reflect on their decision (see Bermúdez, 2009,

chapter 3 for further discussion). Thus, both studies illustrate how looking to the cognitive sciences can inform and criticize ethical theorizing, as the cohesion criterion suggests.

To introduce the next three criteria, I will discuss recent criticisms of virtue ethics (though I intend what I say to apply to other ethical theories). Arguably the most prominent empirical critiques of virtue ethics come from situationists, primarily John Doris (1998, 2002, 2010; see also Gilbert Harman 1999, 2009). These criticisms have focused on the role and robustness of character traits. As Doris (2002) understands the traditional view of character traits within virtue ethics, “If a person possesses a trait, that person will engage in trait-relevant behaviours in trait-relevant eliciting conditions with markedly above chance probability p ” (p. 19). However, Doris and others point to experiments in which the relevant eliciting conditions are present but the traits are either unexpressed or expressed only weakly, due to other, often trivial, situational factors. For example, in Milgram’s (1974) classic study, people were willing to deliver dangerous shocks to another individual when instructed to do so by an experimenter; and in Darley and Batson’s (1973) study, Princeton seminary students were less willing to provide assistance to a person in need of help when they were running late for a lecture. Examples such as these abound, suggesting that character traits are frail and ineffectual.

Situationists also claim that the pervasive causal influence of situational factors suggests that traits, when they are expressed, are only expressed in particular situations. They do not hold across a broad range of conditions but are fragmentary and narrow. So,

for instance, we might express pro-animal virtues when in nature, or when considering what foods to buy, but not when considering whether to donate to the Humane Society or when providing care for our pets.

To state the situationist challenge differently, the traditional view in virtue ethics holds that character traits are enduring dispositions capable of producing behaviors relevant to the respective virtue. They shouldn't function in one setting but not another similar setting, and their efficacy shouldn't vary according to trivial modifications in situation or circumstance. If character traits aren't actually expressed in these ways, then the traditional theory of virtues appears to be problematic. The strategy here is similar to the experiments cited above: virtue ethicists make inaccurate psychological assumptions about character traits, and thus virtue ethics should be modified (or discarded).

There is a further strategy in many of these critiques, however. The frail and fragmentary nature of character traits suggests that being a virtuous agent is difficult, if not impossible. As a result, virtue is *rare*. This is in fact something virtue theorists generally embrace. As Merritt (2000) says:

Now many sympathizers with virtue ethics will want to say, "So what?" The experimental evidence shows only that most people aren't genuinely virtuous. (And haven't we always known this anyway, without needing experimental psychology to reveal it?) That doesn't mean there's a problem with the normative ideal of virtue ethics. It just means that being genuinely virtuous is a rare and difficult achievement. (pp. 367-368)

This strategy on the part of virtue theorists is interesting. Situationists have generally acknowledged that they have no evidence to suggest that virtue is, strictly speaking, impossible, yet they still think its rarity has negative implications for virtue ethics. For instance, Stich and Doris (2005) claim that, at the very least, the burden has shifted:

The advocate of virtue ethics can no longer simply assume that virtue is psychologically possible. If she can't offer compelling evidence—very preferably, more than anecdotal evidence—favouring the claim that virtue is psychologically possible, then she is in the awkward position of forwarding a view that would be undermined if an empirical claim which is not obviously false were to turn out to be true, without offering compelling reason to think that it won't turn out to be true. (p. 121)

This reply emphasizes that virtue ethicists owe an account of how virtue is to be achieved, given its rarity, if not total impossibility.

Thus a further argument can be put to work, in addition to inaccurate psychological assumptions made by virtue ethicists. It is not just that virtue ethicists are wrong about moral psychology; rather, what we know about moral psychology entails that virtue is unachievable, or at least very difficult to achieve. The situationist critique of virtue ethics disputes moral psychological claims, but it also goes beyond these claims to make a further criticism, to the effect that virtue ethics makes demands that cannot be met. This suggests a further criterion of psychological plausibility:

- 2) *Achievability*: An ethical theory is psychologically plausible if it promotes ideas and actions that are achievable by human beings.

I understand this criterion as a projective thesis—whether human beings can meet a theory's demands at some unspecified point in the future. As debates over virtue ethics have shown, this can vary according to the stipulations of the theory. For instance, if only a single person needs to be virtuous to satisfy the demands of virtue theory, then virtue is achievable. Most virtue theorists have more ambitious goals, however, which is why situationist criticisms seem so compelling.

While the achievability criterion is projective, two other criteria serve to link what is achievable in the future to currently existing psychological profiles. The first I

will call psychological *impact*. In essence, this holds that an ethical theory must have an impact on currently existing human agents who make up that theory's primary audience (e.g., some segment of the population). While achievability refers to what might exist at some point in the future, impact refers to the component of a theory that can influence what currently exists. As a criterion of psychological plausibility it is stated as follows:

- 3) *Impact*: An ethical theory is psychologically plausible if the ideas it promotes are capable of having the intended psychological effect on currently existing human beings.

An important feature of this criterion is that impacts must be intended. It would not count, for example, if Kantianism leads me to be a better utilitarian, for that is not the intent of Kantianism.

To illustrate better what I take psychological impact to mean, consider the closely related notion of action guidance. A theory that is action guiding is typically understood to be a theory that offers criteria for right action in particular situations. However, one way a theory can fail to be action guiding is if it does not properly accommodate human psychology. G. A. Cohen (2003) notoriously argued against the idea of action guidance in political philosophy because the aim of political philosophy is not to tell us "what we should do" but rather "what we should think, even when what we should think makes no practical difference" (p. 243). But it would be strange to suggest that what we think is strongly independent from what we do. One common way of determining whether a theory has failed to have the intended psychological impact is to observe someone's behavior. If, for instance, someone in Greene et al.'s trolley experiment claims to adhere to a deontological theory that dictates they follow their

moral duty not to kill, yet they make consequentialist judgments every time (killing one to save five), we will assume that their moral psychology is predominantly consequentialist, and that deontology has failed to have the intended impact.

I do not take the impact criterion to strictly require action guidance, however. What is essential is some specified influence on the psychological profile of currently existing human beings. This will interact with the achievability criterion. Virtue theory, for instance, might be able to have some intended psychological impact on people (e.g., honesty in some specific situation), thus satisfying the impact criterion, but virtue in a more general sense might remain unachievable (e.g., honesty across a broad range of situations is unattainable). This appears to be exactly what situationists aim to show.

The second criterion that links future achievability to currently existing human beings is what I call *transition*. Our moral psychologies are highly plastic, and will change over time, but I do not think this entails the absence of psychological constraints. Rather, if we want to be sure we change in the *right* direction, ethicists should be extra careful in specifying how this transformation will occur.

This requirement is reflected in the quote above from Doris and Stich, in claiming that virtue theorists owe an account of how virtue is to be achieved. Becoming a virtuous agent requires transformation—in many cases extensive transformation. Given this demand, virtue theorists must identify how this process of transformation works. As a criterion of psychological plausibility this can be stated as such:

- 4) *Transition*: If an ethical theory requires currently existing psychological profiles to undergo significant modification, that ethical theory is psychologically plausible only if it can explain how the psychological transition could occur, given current understandings of human psychology.

It is important that this criterion not be interpreted too strictly. *Thorough* understanding of the psychological transition is not necessary to make an ethical theory psychologically plausible. The claim is merely that one should have at least some idea of the attendant psychological demands implied by the theory—to explain how current psychological impacts are connected to future ideals. For example, we can be pretty sure that people are capable of becoming vegetarians, but a theory arguing for vegetarianism is not psychologically plausible if it cannot explain—in at least a general way—how the transition to vegetarianism is supposed to work, psychologically.

I will discuss a third prominent empirical critique, from Jonathan Haidt, in order to introduce the fifth criterion. Haidt's critique focuses on yet another popular theme in the history of ethics: the centrality of harm in moral transgressions. Haidt contends that the sphere of morality goes far beyond harm, to include violations of purity, loyalty, and much else. According to Haidt's research, human morality is determined by six moral foundations, which include liberty/oppression, fairness/cheating, care/harm, loyalty/betrayal, sanctity/degradation and authority/subversion (Haidt & Graham, 2009; Haidt & Kesebir, 2010). Across 12 different world regions, Haidt and his colleagues found significant differences between people who valued fairness and harm on the one hand, and people who valued loyalty, authority, and sanctity on the other (Graham et al., 2011; Graham, Haidt, & Nosek, 2009; Haidt & Joseph, 2004; Van Leeuwen & Park, 2009). This suggests that people possess, at least to a certain extent, *non-overlapping* moral foundations. Some tend to value fairness and harm more than loyalty and dominance hierarchies, while others, though they are not insensitive to harm and justice,

tend to value loyalty and other values associated with dominance hierarchies. This is a problem for ethics because many prominent ethical theories focus primarily on harm and justice, and tend to ignore the other moral foundations.

An example will help illustrate why this is a problem. Although people drawing from any of these foundations will be *attentive* to harm, Haidt and his colleagues found that *moralizing* of non-harmful behaviors was more prevalent among non-Western societies and people who were low in socioeconomic status (Haidt, Koller, & Dias, 1993). Brazilians, to take one example, tended to treat the wearing of uniforms to school as universally binding—a characteristic feature of moral norms—but denied that anyone was harmed by the violation of this norm. A conclusion that can be taken from this research is that a harm-based ethical theory will fail to resonate in the right way, morally, with certain psychological profiles. For example, someone who does not ground their morality in care/harm will be able to perceive and be sensitive to certain types of harm, such as animal cruelty, but without drawing more heavily from care/harm it is unlikely that they will refrain from particular types of wrongdoing *because* they cause harm. They would be less likely to see animal abuse as morally wrong, for instance. Haidt's research thus predicts that ethicists promoting moral principles based in harm are less likely to gain reception among those who ground their moral norms in loyalty and authority (conversely, an ethical theory focused solely on loyalty and tradition would similarly have difficulty in resonating with people who ground morality in harm).

Haidt shares with Doris and Greene the strategy of identifying inaccurate factual assumptions as a way of criticizing ethical claims. Those who think their moral foundations are universal, or think that everyone will recognize the fundamental nature of harm transgressions, are wrong, and presumably this should have implications for the ethical claims such people wish to make (Prinz, 2007). Haidt and Fredrik Bjorklund (2008) press the importance of this when they assert, “All ethical statements should be marked with an asterisk, and the asterisk refers down to a statement of the speaker’s implicit understanding of human nature as it is developed within his culture” (p. 214).

A more fundamental strategy found in Haidt’s work, however, is demonstrating how incompatible moral foundations frustrate the goals of ethical theorizing. If, as Haidt suspects, most ethical theories proposed by professional philosophers appeal primarily to harm, liberty, and fairness, then the results of ethical theorizing will be unpersuasive to large segments of the population, who draw from other moral foundations. Commentators on Haidt’s research have been frustrated by this aspect of his work, as it is not clear if ethical theories do in fact aim to be broadly persuasive. For example, they may only aim for truth, and it may be the case that some moral foundations are more truth-apt than others. *If* persuasion is indeed the goal, however, Haidt’s reasoning seems sound. Constantly making claims only from narrow foundations is bound to be unpersuasive if a significant body of people do not share, or prioritize, those foundations.

As a criterion of psychological plausibility, I will state this strategy as follows:

- 5) *Congruence*: An ethical theory is psychologically plausible if its moral psychological claims and assumptions do not oppose, and are broadly consistent with, the main goals of the theory.

Haidt's research is one particularly prominent application of this strategy. Many ethicists design their theories hoping to convince others who disagree with them, yet, according to Haidt, the content of their theories draw from a small set of moral foundations (e.g., only harm), thus failing to convince others of their moral claims.

Others use this strategy as well, though for goals unrelated to persuasion. Stocker (1976), for instance, provides one classic formulation of this criterion. He argues that a moral theory that does not provide a supporting motivational structure or provides a motivational structure with opposing motivations results in "moral schizophrenia." The basic idea shared by Haidt, Stocker, and others is that an ethical theory becomes psychologically implausible if its moral psychological claims are self-defeating in some way. Consider virtue theory again. If a goal of virtue theory is to make *all* people more virtuous, then it would seem self-defeating to promote rarity. As Doris (1998, p. 512) says, "to what extent does reflection on a few extraordinary individuals facilitate ethically desirable behavior? Or more broadly: what exactly are the practical advantages enjoyed by ideals of virtue?" Focusing only on the moral psychology of those who are already morally excellent would seem to raise difficulties about addressing, psychologically, those who are less excellent. Even if this is false (and I do not wish to take a stance on whether it is), the example helps to illustrate the form of reasoning behind the congruence criterion, and why it is important. Making psychological claims

and assumptions that run counter to one's main goals is presumably something most ethicists would hope to avoid.

There are of course a number of replies one could give in defense of virtue theory (as well as the centrality of harm). As mentioned already, however, for the purposes of my investigation it does not matter if one accepts these particular critiques. What are more important are the general strategies behind these critiques—why the form of critique provided by Greene, Tversky and Kahneman, Doris, and Haidt might be compelling.

More specific criticisms I make of theories in animal ethics in later sections will generally conform to these five criteria of psychological plausibility (which I will henceforth refer to as CPP). That is, I will argue that many ethical theories fail to abide by one or more of these criteria, and I plan to show how and in what ways they do so.

2.1.3 Replacing the PMPR

My argument, then, is that the criteria just outlined are preferable to Flanagan's PMPR. One general reason to prefer the CPP is that they provide much more meaningful constraints on ethical theories than the PMPR. I claimed that the PMPR is too lenient, and fails to place any meaningful constraints on ethical theorizing. For instance, there would seem to be numerous morally commendable psychological profiles that *could* exist under *some* conceivable social arrangement, as the PMPR promotes. It is not clear, however, why such potentialities should be taken seriously if 1) they conflict with everything we currently know about human psychology, 2) it's not clear what social

arrangements would actually be sufficient to make the profiles achievable, 3) they have no psychological impact on currently existing human beings, 4) no ethical theory can explain how such potentialities are to be attained, or 5) they conflict with the explicitly stated goals of specific ethical theories. Ethicists who insist that we retain access to certain psychologically distant profiles must, according to the CPP, address these concerns.

Another benefit provided by my criteria is that they make our ethical options more manageable. The CPP are not intended to be overly restrictive; rather, they crystallize our theoretical options by rejecting those theories that do not meet standards of minimal plausibility. They restrict the range of conceivable options to those that are more likely to have “psychological grip” on ordinary human beings. Certain psychological profiles that we might wish we could have must be rejected, but there will certainly be numerous attractive options remaining. The CPP allow us to focus on these attractive—and more plausible—alternatives.

A third reason to prefer my criteria, complementary to the second, is that they have more to say to currently existing psychological profiles. The CPP offer more specific advice on how to assess theories according to psychological standards, while the PMPR merely says that moral theories should adhere to what is possible for human beings, without specifying what this means. This greater specificity of the CPP has the effect of providing greater assistance to currently existing people in assessing their ethical options. For instance, the PMPR allows for currently existing profiles to undergo nearly limitless change (within “conceivable social arrangements.”). By contrast, the

CPP rule out options for change that are psychologically implausible, and identify steps one could take in making this decision.

A corollary of the two reasons just mentioned is that the CPP use currently existing psychological profiles to limit what ethical theories should promote, if they want to have an impact. This entails that the CPP not only have more to say to currently existing profiles, they also place greater normative significance on current profiles. This offers a different tool for ethical theorizing than does the PMPR. While the PMPR focuses on conceivable social arrangements, and a basic idea of psychological possibility, the CPP place emphasis on currently existing human beings to inform ethical aims. Many ethicists and political theorists have claimed that it counts against an ethical theory if it is not action guiding, or if the only action guidance it can give to agents is to find new action options (e.g., Copp, 2008; Gheaus, 2013; Valentini, 2009; Velleman, 2013). Moral prescriptions are generally expected to offer guidance on options available to the agent, and not just those that exist somewhere else in psychological space. Though the PMPR does not reject action guidance, it does not support it in any direct way. The CPP, by contrast, place emphasis on action guidance by focusing on currently existing profiles, and the moral options currently available to moral agents.

These considerations, as well as those discussed above, provide reasons to accept the CPP over the PMPR as basic constraints on ethical theories. The CPP will be used in later sections to illustrate how animal ethicists in particular can benefit from psychologically plausible theories. Having offered some considerations for basic

psychological plausibility, now I will turn to moderate psychological constraints, or constraints applying to theories that meet the basic conditions outlined thus far.

2.2 Moderate Psychological Constraints

In this section, I will make a distinction between the basic criteria set out above and more moderate constraints on ethical theories. I will do so by elaborating on and clarifying the achievability criterion. The achievability criterion is pertinent here because, as I will explain, many ethicists and political philosophers assume that any psychological limitations can be overcome incrementally, and thus such limitations do not hold any implications for the achievability of ethical theories or moral ideals. I will argue against this idea by turning to two related debates in the philosophical literature: “ought implies can” (OIC) and “ideal theory.” First I discuss the relevance of constraints for political and ethical ideals, primarily as these ideas have been developed in political philosophy. The main problem I identify here is that individuals do not possess the resources (of various sorts) to overcome psychological limitations in the way societies do, and thus individuals’ pursuit of ideals is likely to require significant resource tradeoffs. I then discuss recent debates over the relevance of resource limitations in determining whether various formulations of OIC are true. I argue that when personal resources are limited, as I suggest they often are when attempting to meet ideals, people cannot be expected to meet various ethical prescriptions. This follows from prominent discussions of OIC. I conclude by applying the lessons learned from the OIC literature to debates over ideals.

2.2.1 Achievability and Ideals

An important and illustrative statement of the role of ideals in ethics can be found in Paul Taylor's (1986) classic *Respect for Nature*:

The main function of the ideal is to provide a focus for practical goals. It does this by specifying a kind of world order whose gradual realization is the permanent long-range moral purpose behind the exercise of instrumental rationality by moral agents. In getting a clear grasp of the content of the ideal, agents know the overall direction they wish to take in setting practical goals. The immediate tasks they set for themselves are aimed at changes in the actual world that they believe will make it more closely approximate the ideal world as they conceive of it. It is because they envision the final outcome of their endeavors in terms of the ethical ideal that they use their factual knowledge the way they do in choosing practical ends and the best means to those ends.... The ends and means are sought not for their own sake but for the sake of the ideal...for the sake of making the world a better place by bringing it one step closer to what it should be. (p. 310)

In this passage, Taylor identifies a number of possible benefits provided by ideals, many of which are echoed in the ethical literature on ideals. Ideals provide focus and direction, telling us what to aim for and how—a harmonious conjunction of means and ends. As many other prominent accounts have argued, ideals help us make sense of what we are doing in our moral lives (e.g., Velleman, 2002).⁸ Rawls (1999) famously claimed that ethical theory in fact *presupposes* ideals, because “until the ideal is identified, at least in outline...non-ideal theory lacks an objective, an aim, by reference to which its queries can be answered” (p. 90).

Among ethicists, the richest discussion of achievability comes out of debates over the role of political, social, and economic limitations on ideals. These sorts of

⁸ Aaltola (2012), for instance, claims that ideals “concern goals that will always remain elusive, but without which ethics dwindles into something blank, directionless and cynical.... It is finding the right direction, rather than getting there, that matters” (p. 124).

limitations are almost always considered unproblematic. The unachievability of ideals, on these accounts, is typically understood to be a contingent feature of current political, social, and economic conditions—all of which can be changed.⁹

The strategy usually adopted in overcoming these limitations to ideals is to take slow, incremental steps. This view is shared by just about every commentator on the role of constraints in ethics and political philosophy. For instance, Gilabert and Lawford-Smith (2012) distinguish between hard and soft constraints. Hard constraints are things like the laws of nature, or immutable facts about human nature such as our mortality or need for oxygen, while soft constraints are things that can be modified through learning or altering one's circumstances. Only the former constrain specifically political ideals, because only they are resistant to incremental changes. According to Gilabert and Lawford-Smith, any economic, institutional, and cultural constraints should be considered soft constraints, because “the limits are neither permanent nor absolute” (p. 813).¹⁰

⁹ Farrelly (2007) provides a broader list of relevant constraints, including non-compliance, unfavourable historical, social or economic conditions, indeterminacy, fallibility, disagreement, human vulnerability, human nature, and problems of institutional design. All of these are generally treated as changeable and therefore unproblematic.

¹⁰ Flanagan identifies three different psychological constraints relevant here, ranging from soft to hard on Gilabert and Lawford-Smith's scale. The first is distance in societal circumstances, either historically or culturally. For example, someone born in contemporary America will have a very different view of dogs than will someone born among the Pembans of Tanzania, who believe that God does not like dogs and that spiritual cleansing is required if one should come into contact with a dog. Adopting the Pembani moral outlook on dogs is in some sense a possibility for the contemporary American, but only through radical moral change. A second is the difficulty in changing powerful biases. Transitioning from carnivore to vegetarian, for instance, may require the modification of biases concerning human use of animals, what is considered edible, the importance of cultural traditions, and many others, all of which make some psychological profiles difficult to achieve. The third feature is changing basic

Similarly, Mark Jensen (2009), borrowing from Mele (2006), makes a distinction between three different types of human capacities: abilities we currently express (synchronic), latent abilities that we eventually come to express (diachronic), and latent abilities that we could develop but do not (indirect diachronic). For example, someone could currently be a vegetarian, or if they aren't they could learn to be a vegetarian, or perhaps they never make the transition to vegetarianism but it is within their capacity.¹¹ Jensen argues that most psychological constraints are only relevant to currently expressed abilities, and not diachronic abilities, which are generally the focus of ethical and political ideals. Psychological constraints have no normative force, on his account, because 1) we can usually make changes to our psychological profiles, and 2) most proposed changes are meant to be incremental, thus gradually modifying any constraints.

According to these accounts, psychological limitations merely constitute potential costs or obstacles to change. But these do not alter normative prescriptions. For instance, here are Donaldson and Kymlicka (2011) discussing the obstacles present in converting to vegetarianism:

In the transition period there can be a sense of deprivation of past freedoms and opportunities, and a powerful awareness of the burden of new practices. So there

psychological machinery, such as fundamental psychological processes like reward and punishment. For example, it would be quite difficult, as part of a campaign against meat eating, to modify our ability to find food appetizing or to need to eat food at all.

¹¹ Suppose someone needs to eat meat to survive. If they go more than a few days without meat, they will die. Such an individual would lack both the synchronic and diachronic ability to be a vegetarian; there's no way they could ever learn to become a vegetarian. They could, however, have an indirect diachronic ability to become a vegetarian. For instance, someone could develop medicine to help this individual to survive without meat. Even if this medicine was developed long after the individual has died, the indirect diachronic ability always existed. That person's disposition to be cured by the medicine was always there, such that if the medicine had been discovered sooner, they would have been able to become a vegetarian.

needs to be a transition strategy to deal with this (e.g. incremental changes, lots of experimentation, compensations, and so on). But in judging what are reasonable efforts to bring about the circumstances of justice, the fundamental issue is not the transition costs (which can be offset), but whether the transition leads to fair and sustainable practices in the long term. (p. 202)

This would seem to reduce the significance of any relevant constraints. The difficulty or improbability of realizing an ideal holds no implications for the merit of the ideal. The only real challenge is constructing a transition strategy.¹² And arguing against an ideal on the grounds that it would be difficult to achieve seems defeatist, or as Gilibert and Lawford-Smith say, would be to endorse “cynical realism” about moral progress.¹³

The problem I see with all of these accounts is that they aim at modifying human behavior at a societal level, through the tools available to governments and nations. But ethical prescriptions are frequently made for *individuals*, who have relatively limited resources. Taking incremental steps at a policy or societal level has the benefit of enormous resources and time to complete the change (not to mention that someone else will likely bear the brunt of the costs). Individuals do not have these benefits.¹⁴ Aside

¹² The only question would be whether the cost of achieving the transition is justifiable. And in the case of animals, many ethicists are likely to say yes, they are. Cochrane (2012), for instance, argues that when animal lives are at stake, we should be willing to absorb any costs.

¹³ Rawls (2001) offers a similar analysis, “there is a question about how the limits of the practicable are discerned and what the conditions of our social world in fact are; the problem here is that the limits of the possible are not given by the actual, for we can to a greater or lesser extent change political and social institutions, and much else” (pp. 4–5).

¹⁴ There are also numerous criticisms of ideal theory more broadly. Sen (2006, 2009), for instance, argues that achieving more justice in our current world does not require that we understand what ideal justice is or what it might look like. And Goodin (1995) argues that focusing only on ideals often makes us overlook second- and third-best normative options, where ethical prescriptions might have greater force but still satisfy moral demands. I find these accounts persuasive, but they do not address the criticisms I have in mind. For other good discussions of constraints within the context of ideal and non-ideal theory, see Farrelly (2007), Gheaus (2013), Gilibert (2011), Simmons (2010), Stemplowska (2008), and Swift (2008).

from the obvious differences in financial resources and human capital, individuals also possess limited psychological resources in ways governments and nations do not. Psychologists can measure depletions in ego, willpower, motivation, self-control, and many other psychological features that factor into our moral behavior and limit what we are capable of. These human limitations do not affect governments and nations as much because at a societal level these limitations are more distributed (in a variety of dimensions), and any one government entity can receive support from other governmental entities, all of which have more resources than are available to individual human beings.

As I understand the nature of ideals, meeting them incrementally requires us to pursue a long chain of duties, which will take an extensive period of time, or demand a lot from us over a short period of time, or potentially both.¹⁵ This entails that there will always be significant tradeoffs in meeting any particular set of duties in the service of an ideal. Given the strain on our personal resources, we can only pursue so many ideals, and we pursue some at the expense of others. This exacerbates individuals' limitations in meeting ideals.

These considerations indicate that it makes a difference whether an ethical theory's prescriptions have broader societal support and whether people possess the resources to meet those prescriptions. Donaldson and Kymlicka, for instance, seem to

¹⁵ Brownlee (2010) argues that this is an overly instrumental conception of ideals, and ignores all the ways in which ideals restructure our lives without specifying any particular path or endpoint. I agree with Brownlee, but am focusing specifically on cases where ideals are conceived instrumentally, where the ideal is something that is to be approximated through incremental steps.

assume that individuals can expect infrastructural support for vegetarianism. However, even in wealthy countries (like the U.S.), it is much easier to be a vegetarian in some places than in others. If only minimal support exists, it is less clear whether individuals actually have a duty to convert to vegetarianism, given that they will need to invest extensive personal resources in the conversion, and are likely pursuing other moral ideals that make similar demands on personal resources. For example, someone devoted to famine relief may invest significant time, money, and motivation into famine relief efforts, resources that would thus become unavailable, or sufficiently depleted, such that investing in vegetarianism would be extremely difficult.¹⁶

This raises the broader issue of whether someone can have a reason or an obligation to perform an action they in fact cannot perform, or the claim that is more generally known as “ought implies can” (OIC). Much of this section is relevant to OIC, though the OIC literature tends to neglect psychological “cans.” In the next section, I attempt to carve out a space in the OIC literature for psychological plausibility, before returning to the issues raised here concerning ideals.

2.2.2 “Ought Implies Can” and Psychology

The OIC literature tends to focus solely on physical impossibilities limiting individual acts (Graham, 2011; Vranas, 2007). Common examples and thought-experiments consist of single-shot opportunities, where any psychological limitations are

¹⁶ Moreover, famine relief may require one to support meat production and consumption, depending on where the famine is located and the food options in the region.

due only to lack of willpower (or perhaps competing obligations, creating motivational conflicts). For example, do I have an obligation to stop a murder that is about to happen in Paris, France, when I am currently in College Station, Texas? It would seem not, because I cannot fulfill such an obligation. Do I have an obligation to stop a murder happening in Paris if it were happening next week (assuming I am uniquely capable of preventing the murder)? It would seem I do, and the only limitation would be my ability to will myself through the necessary steps (e.g., buying plane tickets, getting on the plane, etc.). I aim to get a clearer idea of how psychological constraints fit into the OIC literature, particularly with respect to individuals. Specifically, I will highlight certain analogies between these one-shot physical impossibility cases and psychological cases including multiple steps.¹⁷

One recent influential defense of OIC, with implications for psychological plausibility, comes from Streumer (2007). He defends the following claim:

There cannot be a reason for a person to perform an action if it is impossible that this person will perform this action. (p. 351)

The truth of this claim, he argues, demonstrates the truth of a further claim, which is a relatively straightforward formulation of OIC:

It cannot be the case that a person ought to perform an action if this person

¹⁷ One challenge is, as just mentioned, most OIC accounts discuss psychological constraints solely in terms of weakness of will. This matches how ideal theorists tend to discuss psychological limitations as well (e.g., Gilabert & Lawford-Smith, 2012). Weakness of will, though important, encompasses only a small range of the psychological processes relevant to ethics. Another common way of understanding psychological limitations among ideal theorists is as a conflict between selfishness and altruism (e.g., Murphy, 2000). This too includes only a small fraction of the relevant psychological constraints. As such, much of this literature has failed to address the phenomena I see as most pertinent to psychological plausibility.

cannot perform this action. (p. 351)

His main support for these claims comes from a set of thought experiments. I will mention two.

The first he calls the “argument from crazy reasons.” For instance, Jane, a person living in the twenty-first century, seems to have reasons to prevent the crusades, given how bad the crusades were (as well as the slavery and two world wars that followed), even though, to do this, she would have to travel back in time and single-handedly change the course of history. Or suppose Bob, who is standing on the ground, is told that there is a plane full of passengers with failing engines in the sky above him. Bob seems to have a reason to save the passengers on this plane, given how bad plane crashes are, even though, to do this, he would have to jump thousands of feet into the sky and single-handedly repair this plane’s engines.

The second is what Streumer calls the “argument from tables and chairs.” The argument from crazy reasons is supposed to demonstrate that there are clear cases where there is nothing for human agents to do to make a goal more achievable. The argument from tables and chairs generalizes this to indicate a fundamental difference in types of agents: agents who can do things and agents who cannot (or rather, some things are not agents at all). Streumer’s argument is stated as follows:

There cannot be a reason for a table or a chair to perform an action, because it is impossible for a table or a chair to perform an action. When it is impossible for a person to perform an action, this person is in the same position with regard to this action that a table or a chair is in with regard to all actions. Therefore, just as there cannot be a reason for a table or a chair to perform an action, there cannot be a reason for this person to perform this action. (p. 362)

This argument is important because it introduces agential capacities into discussions of

OIC. The agential capacities of human beings are impotent in the argument from crazy reasons, such that agents can take actions (e.g., perhaps Bob could climb the tallest tree) but ultimately can do no more to achieve their goals than could tables and chairs.

Various objections have been raised against Streumer. A popular idea pertaining to ideals is that one has reason to meet an ideal only if one can make a “legitimate attempt” at reaching the ideal, or make a “real try.” Not just any initial goal should count in favor of the plausibility of reaching an ideal—only those that constitute “real tries.” Heuer (2010) understands Streumer’s “argument from crazy reasons” as describing cases where a “real try” is not available; there is nothing for the agent to try to do. An agent can only really try to do something if trying takes him or her closer to achieving that thing. Streumer’s crazy cases do not fulfill this requirement.

Heuer amends Streumer’s claim to accommodate trying, offering the following formulation:

A person has a reason to Φ , even if she cannot Φ , if there is a reason for Φ -ing, and if she can try to Φ (or take other efficient steps).

Heuer further explains what trying consists of:

‘trying to Φ ’...requires that a person takes steps which are appropriate in the sense that they get her closer to Φ -ing: They must be steps in the right direction.¹⁸ (p. 241)

¹⁸ Heuer’s definition of an efficient step is as follows: “A person takes an efficient step towards Φ -ing ... if her action is done with the intention to Φ , and if it is a necessary part of a plan that, if completed, achieves the intended result” (p. 241). Her definition of trying works just as well, however, and is more suited to my discussion.

Streumer (2010) embraces these amendments to his account, though he does not take them to speak against his arguments, as Heuer intends them to. I side with Streumer here, and will show how his position illuminates the relevance of psychological constraints.

Streumer's crazy cases illustrate the difficulty in specifying what it means to take an efficient step towards reaching a goal. There is a wide range of different actions the agents could take without ever making progress towards their goals. For instance, perhaps Bob has time to run to the top of the highest building, cutting the distance between him and the plane dramatically. But even in such a case, Bob isn't actually closer to achieving his goal, and therefore there is no corresponding increase in his obligation to save the plane. Heuer contends that this is problematic only because of Bob's extreme limitations in time and resources. More legitimate and interesting cases, she suggests, are those that include real tries. But this is a mistake: our resources are frequently limited when striving to meet ethical goals, as I argued in the last section. We also frequently take actions towards a goal only to find out that we have not in fact made any progress towards reaching that goal.

Streumer's cases involve extreme examples of physical impossibility to show that ought does indeed imply can. The reason OIC holds in these cases is because of limitations in resources and the resulting inability to identify what would constitute a real try. This same issue arises, I submit, in cases of meeting psychologically difficult ideals. The use of ideals can be thought of as a way to avoid resource limitations and the "real try" problem by providing agents with a series of steps they can meet, over an

extended period of time. The challenge of OIC can be minimized by saying that ought implies “can take the next step” (an idea first proposed by Stocker, 1971). However, Streumer’s crazy cases indicate the difficulties present in taking the next step—we can take a number of steps that do not actually get us closer to our goal. Psychologically difficult ideals are not impossible to meet in the way Streumer’s cases suggest, but similar obstacles do exist. For instance, at every step it will be the case that limitations in personal resources will restrict what people are capable of. When people are so limited, they will be able to take action, but these actions might not constitute real tries.

For illustration, consider Donaldson and Kymlicka’s vegetarian example. I take vegetarianism to possess significant psychological difficulties, at least for some individuals. If individuals do not have broader societal support, they will be in a situation similar to Jane and Bob. They must be able to take efficient steps towards becoming a vegetarian in order to say that they ought to be vegetarians. This follows from Streumer’s account. And given limited resources, there will be times in which this will not be possible. Eating a single meal that excludes meat is obviously insufficient, since presumably everyone has done that at some point. Even if a meatless meal is consumed with the moral ideal in mind, it is simply too meager of an act; it can easily be reversed with the next meal. Even significantly reducing one’s meat consumption might not count either, depending on broader societal support. If reduced meat consumption is unsustainable, for whatever reason, and must eventually be reversed, it is not clear why we would count a reduction in meat consumption for even a considerable length of time—say, a year—as a step towards becoming a vegetarian. If resources are sufficiently

depleted such that any step the aspiring vegetarian might take will not necessarily bring them any closer to becoming a vegetarian, then there would be no obligation to become a vegetarian. It would not be, strictly speaking, impossible, but the limitations in resources would still dictate that no moral obligation exists.

To briefly consider another example, let's return to the previous discussion of situationist critiques of virtue theory. Suppose a virtue theorist recommends I act as a virtuous agent would. Situationists would contend, in opposition, that I am constrained by my inability to overcome situational constraints. The virtue theorist could outline a series of steps for me to meet, but this would just delude me into thinking that situational constraints could be overcome. This is analogous to what is going on in Streumer's cases. Human agents attempting to be more virtuous can take a number of actions that are attributable to the ideal of acting like a virtuous agent, but these "don't count" unless they constitute real tries. Situationist critiques suggest that many attempts to be more virtuous do not meet this standard.

To briefly summarize, my main claim is that when resources are limited, people cannot be expected to meet psychologically difficult ideals. The resource limitations pertinent to OIC are also pertinent to meeting ideals. Ethicists tend to think of OIC in terms of physical impossibilities, but, I have suggested, there are clear parallels to psychological impossibilities as well as psychological difficulties more generally. The final section will elaborate on the implications of this and address certain objections.

2.3 Conclusion: Implications of Psychologically Difficult Ideals

I will conclude by discussing two possible objections someone might have to my arguments. These are not exhaustive, but are prominent in the literature on ideals and OIC. I will also discuss one example from the animal ethics literature to illustrate the implications of my argument in this section.

A natural objection someone might have to the argument in the previous section is that personal resources are not always so limited. And even if they are, we usually will not know when. Various accounts in the OIC literature have acknowledged this possibility by arguing that ought does not “conversationally” imply can. That is, in moral deliberation we often disregard considerations of what we are capable of in order to more directly assess what we ought to do. These accounts attempt to retain discussions of moral obligations in moral discourse, even if occasionally we cannot meet such obligations.

Sinnott-Armstrong (1984), for example, accepts the truth of OIC, but also argues for the claim “if cannot, might yet be true that I ought.” This claim allows for changes in circumstances. For example, perhaps I currently have no obligation to become a vegetarian because I have no societal support, but if this support improved, I would need to possess the ideal of vegetarianism in order to take advantage of it. Or perhaps I cannot take efficient steps towards an ideal right now, but I need to keep the ideal in mind for the future when such steps become available.¹⁹

¹⁹ A related objection comes from Estlund (2011), who argues that when people possess limited resources, they have a responsibility to increase their access to resources, such that actions they

I do not intend to deny these possibilities. However, there must be some standard for determining which ideals should be pursued. There are likely to be many competing ideals with significant ethical merit. We possess many oughts conversationally. Appealing to the difficulty of ideals is one way of choosing among these oughts, and eliminating some from consideration. Moral ideals must be carefully selected to take into account various constraints, and help individuals either overcome or work within them. Referring only to what obligations would hold across counterfactual situations fails to address currently existing constraints. While I agree that ideals should be retained in moral discourse, some ideals hold greater merit in moral discourse by virtue of their plausibility. Those that address the most plausible “cans” are more helpful than those that address only future possibilities.

Another objection comes from Gilabert and Lawford-Smith’s remarks above, claiming that using psychological limitations in the way I have is a form of “cynical realism.” Am I not being unfair to the flexibility of individuals and the potential for things to change? Why should we think that psychological change is anything like Streumer’s crazy cases?

Consider the following spectrum of positions one could take on the normative implications of psychologically improbable moral paradigms:

Cynical Realism: individuals can never change and there is no use asking them to try.
Optimistic Realism: individuals probably will not change, but that should not stop us from asking them to try.

cannot currently perform will be within their capacity in the future. However, this ignores the fact that resource limitations also constrain people’s pursuit of resource accessibility, and so I will not discuss this objection in any detail.

Practical Idealism: individuals can change, but we should focus on the most realistic options when asking them to try.

Starry-Eyed Idealism: individuals can change; it just takes time and effort.

I reject both of the extremes here, and favor a position that falls somewhere in between Optimistic Realism and Practical Idealism. I am certain that individuals can and will change, but not without significant support. And without this support, they cannot meet various moral obligations, and thus do not have those obligations. Individuals cannot count on social and cultural change to actually occur, and they cannot enact such change on their own. If such a state of affairs exists, it would be particularly Starry-Eyed to maintain that individuals' moral obligations remain unchanged. Saying this, however, does not make me a Cynical Realist. Rather, I am nudging ethicists toward the huge unexplored conceptual space between these extremes.

I will discuss one example from the animal ethics literature to illustrate what I have in mind here. Gary Steiner (2013, pp. 146-147) argues that some of the most prominent theories in animal ethics are not sufficiently demanding (also see Francione 2000). Animal welfare positions, he argues, make too many compromises. If they were consistent, they would endorse the animal rights position he favors. Animal welfarists think it is permissible to kill and eat animals, so long as they are not made to suffer in the process, while animal rightists think all human use of animals is impermissible. In making concessions to meat-eating, animal welfarists fool us into thinking that we are making moral progress, when in fact things have not improved much at all.

Steiner's position on animal rights and animal welfare can best be characterized as an instance of Starry-Eyed Idealism. The animal rights position claims that it does not

matter what people currently believe or do, killing animals is wrong, and people can eventually come to realize this, and so that is what we should demand of human agents. In arguing against animal welfarists, Steiner is rejecting something along the lines of Practical Idealism. Though I don't wish to make an argument for Practical Idealism specifically, it is much more psychologically plausible than Steiner's animal rights position. It is notable that animal welfarism still demands quite a lot from human agents and improves treatment of animals considerably. Whether it somehow reinforces poor treatment of animals is an empirical question. For his Starry-Eyed Idealism, Steiner comes off as a Cynical Realist in assuming that animal welfarism will be the last stop for improving animal treatment.²⁰

This should serve as an example for animal ethicists who wish to make their theories more psychologically plausible. Even with the constraints I have identified—which might seem overly strict to some—there still exists substantial room for moral progress. Though I intend for my criteria to have a significant impact on animal ethics, this does not imply that animal ethicists should settle for poor treatment of animals. In fact, if my arguments have been right, ethicists contribute to poor treatment of animals by proposing ineffective and psychologically implausible ideals. Ethicists can improve their theories, however, by becoming more informed about the relevant psychological issues and working to meet the criteria I have outlined.

²⁰ Garner (2013) makes similar criticisms of Steiner and Francione in the context of nonideal theories of justice. Garner does not discuss the issue of psychological constraints, however, and my project is not concerned with justice specifically, so Garner's views will not be discussed further.

3. MENTALIZING ANIMALS: IMPLICATIONS FOR ANIMAL ETHICS

Ethicists have tended to treat the *psychology* of attributing mental states to animals (henceforth “mentalizing”) as an entirely separate issue from the *moral importance* of animals’ mental states. In this section I bring these two issues together. I draw from recent empirical research on mentalizing to argue that ordinary human agents use what are generally called *phenomenal* mental states to assign moral considerability to animals, or to identify which animals are owed basic moral obligations. For normal psychological profiles, I argue, phenomenal mental states are both necessary and sufficient for judging an animal to be morally considerable. This claim is central to what I will refer to as the *phenomenal account* of mentalizing.

While the phenomenal account is congruent in many ways with current consensus in animal ethics, I also argue that the behavioral triggers we possess for attributing phenomenal states to animals are quite narrow—primarily animals that look and act like human beings. This presents a challenge to animal ethics, as most theories require people to recognize that a wide range of animals, far beyond those that look and act like human beings, are capable of phenomenal experiences (also usually described as *sentience*).

I will frame the implications of the phenomenal account for animal ethics in terms of moral considerability, as explained above, and moral significance, which refers to animals’ relative moral importance (e.g., a frog compared to a human being). In the first section, I explain the distinction between considerability and significance in more detail in order to sketch, in a general way, the relationship between animal ethics and the

psychology of mentalizing. I will outline the most fundamental ways animal ethicists use mental states to assign moral importance and summarize how this compares to research on mentalizing.

The second section provides empirical evidence in favor of the phenomenal account, including evidence that we attribute phenomenal states based on an entity's physical and functional similarity to human beings. In the literature, this evidence is framed in terms of Experience, which is synonymous with phenomenal states, and includes the ability to feel pain, fear, joy, and other emotions, and in terms of Agency, which includes the ability to use language, form intentions, plan, and make decisions. I argue that the evidence indicates that the perception of Experiential states are necessary and sufficient, psychologically, for assigning moral considerability. I also address potential objections to my argument, specifically those claiming that Agency is necessary and sufficient for assigning moral considerability.

The third section explores the relevance of the phenomenal account for moral significance. I highlight the fact that limitations in judgments about moral considerability necessarily limit judgments about moral significance as well. Limitations in phenomenal mentalizing narrow the range of animals deemed morally considerable, which, in turn, prohibits animals' agential states from ever becoming relevant.

In the final section, I assess the psychological plausibility of theories in animal ethics in accordance with the criteria outlined in the previous section. My overall conclusion is that purely agential theories are so psychologically implausible that they should be abandoned, but that many other views in animal ethics are sufficiently

congruent with the phenomenal account to only require certain revisions. The phenomenal account challenges animal ethics in various ways, but it also suggests that the field has been moving in a positive direction for the last 40 years.

3.1 Moral Considerability and Moral Significance

In this section, I first explain how animal ethicists understand and use the concepts of moral considerability and moral significance. I then discuss how ethicists use mental states to assign moral considerability and significance. I will also briefly sketch the phenomenal account in order to summarize how ethicists' use of animal mental states compares to research on the psychology of mentalizing.

3.1.1 The Basic Distinction

A common distinction in animal ethics, classically made by Kenneth Goodpaster (1978), is between moral considerability and moral significance. Moral considerability refers to an entity's status as being worthy of basic moral attention—as deserving moral attention at all—while moral significance refers to an entity's relative moral importance (e.g., the moral status of a human being as opposed to a frog). Being morally considerable marks off an agent as having surpassed a *minimum* threshold, beyond which is the realm of entities to whom we possess varying degrees of moral responsibilities. There are many additional ways of carving up our moral judgments concerning animals, but these two concepts capture the predominant approach within animal ethics. I will discuss each in turn.

Other phrases used as alternatives for moral considerability are “moral status” and “moral standing.” In this section, I take these terms to be synonymous.²¹ Arguably the most common way of thinking about moral considerability is in terms of an entity’s interests. The interests of a morally considerable entity *matter*, ethically. As David DeGrazia (2008) says “To say that X has moral status is to say that (1) moral agents have obligations regarding X, (2) X has interests, and (3) the obligations are based on X’s interests” (p. 183). For instance, many ethicists take there to be a conceptual link between suffering and aversiveness, such that the suffering of an animal entails that it has an interest in avoiding that suffering. This automatically confers on us a *prima facie* moral obligation to prevent or alleviate that suffering.

The chief question, which forms the bulk of discussion among animal ethicists, is what being morally considerable implies, or what kind of treatment should follow from being morally considerable. For instance, perhaps this only means an animal’s interests should not be hindered, or perhaps instead it means its interests should be furthered; the animal should be assisted in meeting those interests. In some cases it is not clear if either is the case (e.g. wild animals who, it is often argued, should not be interfered with). As was discussed in the last section, some ethicists argue that killing and eating animals is consistent with their moral considerability, so long as they are not made to suffer. Conversely, others argue that being morally considerable means that animals should not

²¹ Mary Anne Warren’s (2000) classic definition also seems to take these three terms as ultimately interchangeable: “to have moral status is to be morally considerable, or to have moral standing. It is to be an entity towards which moral agents have, or can have, moral obligations” (p. 3).

be used for human ends, regardless of whether or not that use would cause suffering.

These latter issues raise questions pertaining to moral significance. Even if animals are morally considerable, one might think that human interest in eating them is justifiable because of our relative moral importance—our interests trump theirs. Their moral considerability still dictates that they should not be made to suffer, but this might be all they are entitled to. If an animal is entitled to much more than this, it is because we have assigned them greater moral significance. For example, it is often thought that the great apes cannot be eaten because their moral significance is on par with our own (indeed, human beings are classified as one of the great apes).

The feature animal ethicists use most often to assign moral considerability and moral significance is an animal's psychological capacities. Apes' moral significance, for instance, is thought to be a function of their intelligence. The next section will explain how mental states are used to dictate moral importance.

3.1.2 Moral Importance and Mental States

As mentioned above, animals' capacity for suffering is widely understood to provide grounds for assigning moral considerability. However, suffering is just one of many mental states thought to indicate *sentience*, which provides a more broad-based justification for assigning moral considerability. As Peter Singer defines it in *Animal Liberation*, sentience is “the capacity to suffer and/or experience enjoyment” (1990, p. 8). However, as Singer goes on to explain, sentience consists of mental states generally characterized as *phenomenally conscious*. This means that an animal's suffering and joy

have a subjective quality and feel a certain way to that animal.²² Varner (2012) thus defines sentience more aptly as “the capacity for phenomenally conscious suffering and/or enjoyment” (p. 108). This definition will be important throughout this section.

A potential objection to this definition must be considered before proceeding. Sytsma and Machery (2010) object to the use of the term “phenomenal consciousness” because of the debate within consciousness studies over whether phenomenality is indeed central to consciousness. According to them, characterizing consciousness as “phenomenal” begs the question against theories of consciousness that deny that conscious states are characterized by their subjective quality. I do not wish to take a stance on this issue, as the problem can be sidestepped in the case of animals by focusing on emotional states. Emotions are often said to be *valenced*, which means they possess a positive or negative subjective quality (or possess positive or negative hedonic value). This is indeed why Singer and many other animal ethicists have focused on suffering and joy—they are clearly valenced. Thus, valenced states are unambiguously phenomenal. Sytsma and Machery’s point is that valenced states do not exhaust the full

²² For instance, in *Practical Ethics* (2012) Singer says, “Terms like ‘pleasure’ and ‘happiness’ lack precision, but it is clear that they refer to something that is experienced or felt—in other words, to states of consciousness” (2012, p. 77). He furthermore connects these felt experiences to moral considerability, “If a being is not capable of suffering, or of experiencing enjoyment or happiness, there is nothing to be taken into account” (p. 50). Mark Bernstein, in his book *On Moral Considerability* (1998), similarly asserts, “experientialism dictates that all and only those with moral standing are phenomenological or sentient individuals” (p. 24). The well-known animal welfare scientist Francoise Wemelsfelder (1999) also seems to support this basic idea, “The concept of consciousness, as it functions in common-sense interaction with animals, denotes that animals are not mere objects but subjects; that is, it indicates that a level of behavioural organization is present which requires a non-mechanistic, subject-related, first person perspective level of explanation” (p. 42). This connection between conscious states and moral considerability is nicely drawn together in the phrase “experiential well-being” (Palmer, 2010).

range of conscious mental states. This point will not affect the discussion here, however, nor does it affect the definition of sentience.

There are three broad views one can take on the relationship between animals' mental states and moral importance. The first I will call *pure phenomenal* views. These use animals' capacity for phenomenal experiences to grant them great moral significance. Steiner (2008, 2013) and Francione's (2000) animal rights theories typify this view. As discussed in the previous section, they argue that animals' sentience entails that they cannot be used for human ends. This does not mean animals and human beings possess equal moral significance, but the gap has been significantly narrowed, based solely on phenomenal states.

Pure phenomenal views, as I understand them, are polar opposites of the second type of view, which I will call *pure agential* views. As mentioned above, common agential capacities include the ability to use language, form intentions, plan, and make decisions. Pure phenomenal views exclude the importance of agential abilities, while pure agential views exclude the importance of phenomenal abilities. Raymond Frey and Donald Davidson's accounts of language typify this view. For both Frey (as expressed in his earliest work) and Davidson, language is a prerequisite for even moral considerability. Frey (1980) argues that language is necessary for beliefs, beliefs are necessary for desires, and desires are necessary for interests. He reasons that an agent must believe he or she has a deficiency with respect to some need of theirs (e.g., water) in order to desire that need. So although animals can respond to pain in such a way that indicates they need to avoid pain, they cannot truly desire to avoid that pain without a

relevant belief. On Frey's account of belief, having a belief consists of judging sentences to be true or false, which is a linguistic activity. Since animals do not have language, they cannot have beliefs, desires, or interests, thereby excluding them from moral considerability.²³

Davidson holds a similar view. As he says, "a creature cannot have a thought unless it has a language" (1985, p. 477). On Davidson's holistic view of language, in order to have a belief, one must have many general beliefs, and this requires a network of communicators. Animals, he argues, do not have this. For example, perhaps a dog understands, in some sense, that his owner is home, but he does not know that his owner is called Mr. Smith, is the president of a bank, and a number of other facts that ordinary language users would know (Davidson, 1980, p. 164). This limits the sorts of beliefs we can ascribe to the dog. As Davidson (1985) explains, "We identify thoughts, distinguish between them, describe them for what they are, only as they can be located within a dense network of related beliefs. If we really can intelligibly ascribe single beliefs to a dog, we must be able to imagine how we would decide whether the dog has many other beliefs of the kind necessary for making sense of the first" (p. 475).

Pure agential views, like pure phenomenal views, narrow the gap between moral considerability and significance. Frey and Davidson's account of language entails that

²³ Frey (2014) clearly states that he thinks animals do have beliefs and desires and can suffer, indicating that his views have changed. His discussion of agential abilities has also shifted to primarily emphasize the differences in autonomy between humans and animals, rather than language (Frey, 1987, 2014). What I say about Frey here and throughout the dissertation should thus be understood as pertaining primarily to his early views. I adopt this strategy because his early views were widely influential, are illustrative of the pure agential position, and are similar in important respects to Davidson's.

animals are not morally considerable, and so lack any moral significance. However, if an animal turns out to be a language-user, Frey and Davidson's view would automatically grant that animal great moral significance, regardless of any other capacities it might possess.²⁴ Great apes, for example, would not need to have the capacity for sentience so long as they are highly linguistic—they would still be on moral par with human beings.

The third type of view I will simply call *mixed*. Theories in animal ethics can be understood on a spectrum from pure phenomenality to pure agency, the majority of which fall in between these two poles, and are thus mixed. I will discuss Varner's (2012) categories of sentient, near-persons, and persons to illustrate how mixed views generally work.

According to Varner's (2012, p. 113, 123) well-known taxonomy of pain-feeling animals, the animals that appear to be sentient (in the phenomenal sense) include all vertebrates, which encompass all mammals, birds, reptiles, and amphibians. Thus, all vertebrates should be judged morally considerable, and different levels of moral significance can be granted among the different types of vertebrates, some of which are persons or near-persons. Persons and near-persons, on Varner's account, possess an array of cognitive capacities that qualify them for great moral significance: near-persons can think about their past and their future; persons possess a biographical sense of self

²⁴ Carruthers (1992) and Korsgaard (1996) seem to hold views similar to the early Frey and Davidson in that rationality, according to them, is a prerequisite for moral standing (though their conceptions of rationality diverge from one another). They differ from the early Frey and Davidson, however, in that they grant beliefs, desires, and some level of intelligence to animals. Since the early Frey and Davidson are "purer" in this respect, I will not discuss Carruthers and Korsgaard's views in any detail. Suffice to say that they are on the agential end of the spectrum.

and the ability to see their lives in narrative terms. According to Varner, only human beings are persons, while apes, cetaceans (e.g., dolphins), corvids (e.g., Western Scrub Jays), and elephants are the best candidates for near-persons, based on the most current empirical research (which he reviews in detail). Importantly, the reason these cognitive capacities matter, on Varner's account, is that they enhance an animal's phenomenal states. As he explains the importance of narrative abilities, "the abilities to consciously remember the past and to consciously anticipate the future allow the individual to reexperience good (and bad) states of consciousness and to anticipate (and dread) future experiences" (p. 162). Narratives are phenomenally rich, and the more phenomenally rich they are, the more moral significance they carry. Meeting one's goals, for instance, can be thought of as realizing—consciously—that one's desires have been satisfied. The desires (or goals) provide one layer of phenomenality, and the conscious realization of those desires (or goals) being met or frustrated adds another layer, and so on.

Varner's mixed approach maintains that some animals are more morally significant because of the way cognitive capacities affect their phenomenal states. This places his view closer to the phenomenal end of the spectrum. However, arguably a more common view is to focus on agential states, and exclude phenomenal states, when discussing moral significance. This is illustrated by Frey and Davidson's views discussed above. Agential states enhance an entity's moral significance without entailing any change in phenomenality. We can thus make a further distinction among mixed views, between those that emphasize agency and those that emphasize phenomenal states. For example, suppose Frey and Davidson modify their view to allow that animals

can suffer (phenomenally), and are thus morally considerable, but they maintain a purely agential account of language. Such a modification, though it introduces phenomenality at one level, might still allow humans to use all animals in a wide variety of exploitative ways (so long as they do not cause suffering) because humans are language-users and animals are not.

3.1.3 Animal Ethics and the Psychology of Mentalizing

In the next section I will present research on the psychology of mentalizing in support of the phenomenal account. First, however, I will outline the incongruences I see between the theoretical options provided above and the psychology of mentalizing. In order to do this, I will briefly sketch the phenomenal account.

The phenomenal account holds that ordinary human agents use phenomenal states to assign moral considerability to animals. More precisely, according to human psychology, phenomenal mental states are both necessary and sufficient for judging an animal to be morally considerable. The moral psychological claim made by the phenomenal account is captured nicely by Philip Robbins (2008):

If a being has phenomenal experience of any sort (i.e., if there is something that it is like to be that being), we are morally obliged to take that experience into consideration. In other words, it is not just that we feel morally obliged to consider others' pains; we also feel morally obliged to consider their pleasures, and perhaps even their hedonically neutral experiences, as factors in deciding how we should act. (p. 20)

The basic idea is that something about an animal's ability to experience phenomenal states (of which pain is a perspicuous example) naturally leads people to think it is worthy of the most basic form of moral consideration.

The connection the phenomenal account makes between phenomenal states and moral considerability matches well with current consensus in animal ethics. Though ethicists dispute what a phenomenal being is owed, they are broadly in agreement that such a being is owed *something*. This is what distinguishes many animals from various other nonhuman entities (e.g. trees).

However, this does not necessarily mean the phenomenal account favors pure phenomenal views. The phenomenal account holds that an animal will be assigned moral considerability if it possesses phenomenal states. This entails that the animal is morally significant to some degree—it has reached the basic level of significance—but it does not hold any further implications for the animal's moral significance. As a psychological thesis, the phenomenal account does not claim that attributing phenomenal states will lead people to assign great moral importance to an animal. It only holds that phenomenal states determine basic judgments of moral considerability. The phenomenal account does not speak against pure phenomenal views in ethics, but it does not directly support them either, except insofar as pure phenomenal views hold phenomenal states to be crucial for moral considerability.

The phenomenal account does, however, speak against pure agency views. Frey and Davidson's linguistic argument, for example, is particularly problematic. According to the phenomenal account, ordinary human beings are likely to reject this approach

because it does not grant any role to phenomenal states.²⁵ People perceive the ability to have linguistically-based beliefs to be separate from the question of animals' moral considerability. This has implications for moral significance as well. According to the phenomenal account, an entity thought to be highly linguistic but lacking phenomenal experiences would not even be morally considerable. And an entity that is not morally considerable has no moral significance. Robots, for instance, as indicated in the research discussed below, possess various complex cognitive abilities, and can simulate certain aspects of language, but are often denied moral considerability (and thereby moral significance) because these abilities do not entail any corresponding phenomenal states.

The general implications of the phenomenal account, then, are that pure phenomenal views of moral considerability are supported, but in a somewhat trivial way, while pure agential views are entirely incongruent with human psychology. Similar implications hold for mixed views. Varner's view, for example, appears to be strongly supported by the phenomenal account, because agential states are important only insofar as they modify phenomenal states. Views that take cognitive capacities to be important in their own right, however—regardless of how they relate to phenomenality—are not congruent with the phenomenal account. However, the exact implications for mixed views will vary according to the specifics of the theory. I will return to this issue in the final section, after examining the evidence in favor of the phenomenal account in more detail.

²⁵ It should be emphasized that Frey in particular does think that an entity's experiential life is what matters for determining its moral status. The problem, in both his earlier and his later views, is that he does not think there is much experience present without language.

3.2 The Phenomenal Account

Here I review evidence in favor of the phenomenal account. I also consider evidence that might be taken to oppose the phenomenal account. Specifically, I will review evidence for the necessity and sufficiency of agential abilities in determining moral considerability. I argue that agency is not necessary for moral considerability and is sufficient *only* when it entails a change in phenomenal states. Lastly, I review evidence indicating that we attribute phenomenal states to entities according to their physical and functional similarity to human beings, which presents a challenge for animal ethics.

The label “phenomenal account,” at least with respect to the link between mentalizing and moral psychology, originated with Philip Robbins and Anthony Jack (2006), but the relevant research has been discussed primarily in the context of the “valence account,” as proposed by Jesse Prinz, Joshua Knobe, Justin Sytsma, and Edouard Machery. The valence account is a more specific variety of phenomenal account, focusing particularly on the role of emotion, or hedonic mental states, in determining moral judgments. The evidence provided by each account is quite similar, and both accounts make similar hypotheses concerning the relationship between mentalizing and moral judgments about animals. Thus I will discuss the relevant evidence under the single heading of the phenomenal account.

3.2.1 Phenomenal States and Moral Considerability

The departure point for discussing the connection between phenomenal mental

states and moral judgments is Gray, Gray, and Wegner's (2007) influential two-dimensional account of mind perception. They conducted a relatively large survey (over 2000 respondents) in which people were asked to compare different types of agents on a wide variety of mental attributes. Participants in the survey were presented with pictures and descriptions of 13 different agents²⁶ as well as descriptions of 18 different mental states.²⁷ The agents and mental states were then presented in 78 pairwise comparisons. For instance, a chimpanzee and a human fetus might be presented as a pair, and people would have to rate which one was more capable of experiencing pain (or if they were equal). Gray et al. used these responses to create a mind index, categorized according to two dimensions. One categorizing feature they termed Agency, which referred to an agent's ability to make choices, have intentions, and control their own thoughts. Entities that scored high on Agency included God and human beings. The other categorizing feature was Experience, which referred to an agent's ability to have sensations and feel emotions such as hunger, fear, pain, and pleasure. Entities that scored high on Experience included chimpanzees and human infants.

As mentioned above, experiential mental states can be understood as synonymous with phenomenal mental states. Indeed, many cite this experiment to argue that the possession of phenomenal mental states determines mentalizing specifically

²⁶ These included a frog, a dog, a young chimpanzee, a human fetus, a human infant, a 5-year-old human female, a 38-year-old human female, a 30-year-old human male, the person taking the survey (a self-rating), a man in a persistent vegetative state, a deceased woman (described so as to elicit opinions on deceased people in general), God, an intelligent robot.

²⁷ These included communication, consciousness, desire, embarrassment, emotion recognition, fear, hunger, joy, memory, morality, pain, personality, planning, pleasure, pride, rage, self control, thought.

related to morality. The primary piece of evidence Gray et al. cite is from a single question they posed to participants, asking which entities they would avoid inflicting harm upon.²⁸ They found that the desire to avoid harming an entity correlated significantly more strongly with Experience than with Agency ($r=0.85$ versus $r=0.26$).

The phenomenal account generalizes Gray et al.'s data to claim that moral judgments will diverge as a function of the two dimensions identified. Put simply, the phenomenal account predicts that agents who are attributed experiential mental states will be judged morally considerable. I will defend a stronger claim: being attributed experiential states is both necessary and sufficient for being judged morally considerable. This claim will be defended and elaborated upon once the relevant evidence is in full view.

Studies conducted by Justin Sytsma and Edouard Machery (2010, 2012) provide crucial support for the phenomenal account. Their surveys focused on mentalizing not animals but robots. They found that when people were asked to attribute mental states to a robot performing simple tasks, they were willing to attribute to the robot the ability to perceive the color red and smell an unknown chemical substance. They were not willing, however, to attribute the mental states of anger, pain, or the ability to smell a banana. Participants were willing to attribute all of these states when the agent in question was a human being and not a robot. The explanation for these results offered by Sytsma and Machery is that unknown chemicals and the color red do not produce valenced mental

²⁸ The exact question asked was "If you were forced to harm one of these characters, which one would it be more painful for you to harm?"

states—they have no hedonic value and do not feel any certain way when experienced—and so are considered within the capacities of robot perception. These results illustrate that mentalizing different types of agents varies as a function of whether those agents are perceived to experience the positive or negative feeling of certain mental states.²⁹

Another important piece of evidence for the phenomenal account comes from Bastian, Laham, Wilson, Haslam, and Koval (2011). They presented participants with descriptions of 24 groups of people that varied on dimensions of Experience and Agency (groups were relatively broad, including athletes, children, lawyers, and so on). For example, those high in Experience were described as “emotionally responsive” and “warm towards others.” Those high in Agency were described as “culturally refined” and “rational or logical.” They then presented participants with a scale from Gray and Wegner (2009) designed to measure experiential states (including an expanded set of questions querying group members’ ability to feel pain). They also asked participants whether they would “intervene” or “take a moral stand” for members of each group, if they were recipients of moral harm (e.g., if someone broke a promise to them). As predicted, the experiential descriptions were positively correlated with willingness to take moral action. Agential descriptions were in fact slightly negatively correlated with willingness to take moral action.

²⁹ Buckwalter & Phelan (2013) provide evidence to suggest that these mental state attributions to the robot vary according to differences in functional abilities. They found that people would attribute valenced states if the robot was described as being functionally capable of doing so. Though this provides evidence against Sytsma and Machery (2010), it is consistent with how I understand the role of behaviors in phenomenal mentalizing. More will be explained on this below.

The two studies just mentioned provide general support for the phenomenal account. Now I will mention a series of studies that provide support for the specific role of phenomenal states in moral judgments concerning animals. One of the first pieces of evidence for the phenomenal account, with respect to animals, comes from surveys conducted by Joshua Knobe and Jesse Prinz (2008). They presented people with descriptions of a researcher studying fish. In one condition, the researcher was described as studying fishes' ability to remember where food is in a lake. In the other condition, the researcher was studying fish's ability to have feelings. The participants were then asked, "Why do you think he [the researcher] might want to know this? Why might the question be important to him?" The results were quite striking: 100 percent of the participants in the feeling group connected the study to moral issues,³⁰ while 100 percent of the participants in the memory condition mentioned items related to predicting and explaining the behavior of fish³¹ (with only 9 percent mentioning moral issues). This study is somewhat crude, but the stark difference in results for the phenomenal question and the agential question are nonetheless significant (one could argue the crudeness makes the results even more surprising).

More recently, Jack and Robbins (2012) and Sytsma and Machery (2012) have provided evidence to support the phenomenal account. Jack and Robbins asked people to

³⁰ Sample response: "He might want to know whether fish genuinely feel things because in doing his job, he does lots of things to the fish that might possibly hurt them if they can really feel things. It might be important to him to find out if he causes them pain because he might feel it is unethical or immoral to cause harm to other things. He could hold this belief for several reasons such as religion."

³¹ Sample response: "So it will be easier to feed them, b/c he only has to distribute food in one place or so he'll know where to go in order to give bait, if they are capable of remembering such things."

read stories about the harvesting of lobsters. In one condition people were told that lobsters possessed the sorts of states that would be considered Agential in Gray et al.'s sense: they were described as intelligent, able to perform elaborate foraging strategies, and having great memories. In this condition lobsters were also described as feeling little to no emotion. In a second condition participants were asked about states that were consistent with Experience in Gray et al.'s study: lobsters were described as possessing the ability to feel emotions such as depression and anxiety but not having much intelligence. Participants were asked to rate on a 10-point scale how concerned they were about lobsters, how they would feel if they themselves were harvesting the lobsters, and how severe the penalty should be if the harvesting was made illegal. In all three cases, those in the Experience condition scored significantly higher than those in the Agency condition.

Sytsma and Machery conducted a study similar to this one but focused on primates. They presented people with a story in which monkeys were being used to test the effects of wound-healing antibiotics and scientists were puzzling over which of five species to use for the experiment. They again varied both Experience (described as the ability feel pleasure and pain) and Agency (described as the capacity for intelligence and inquisitiveness) in describing the species under question. Participants in all conditions were asked to rate on a 7-point scale whether it was morally wrong for the scientists to use the particular species presented to them. They found that Experience had a significant impact on moral ratings but Agency did not.

The evidence reviewed thus far indicates that attributing phenomenal states to

animals is strongly correlated with judgments of moral considerability.³² It also seems to indicate that agential states are not involved with such judgments. However, I also made the stronger argument that the phenomenal account holds that phenomenal states are both necessary and sufficient for moral considerability. To substantiate this, I need to further show the insignificance of agential states in determining judgments about moral considerability.

3.2.2 Agential States and Moral Considerability

Sytsma and Machery (2012) argue that “Experience and Agency are important, independent cues for lay ascriptions of moral standing” (p. 11). As outlined above, they found that describing monkeys in experiential terms was more likely to produce moral concern than describing them in agential terms. However, the influence of agential terms was very close to statistical significance ($p=0.13$). The average ratings of moral concern (on a 7 point scale) for agency were also relatively high (4.37 for high agency and low experience; by comparison, high experience and low agency was 5.34). This provides some evidence in favor of the role of agency in determining judgments of moral considerability.

Their main source of evidence for the role of agency, however, comes from

³² One other suggestive, but nonessential, piece of evidence is that mental state ascriptions to animals seem to utilize the same emotion processing areas of the brain that are used to ascribe mental states to human beings. For instance, in one study, Mitchell, Banaji, and Macrae (2005) asked people to judge whether certain words could “describe a person” or “describe a dog” (e.g., the word “curious”). When mental predicates were judged to be applicable to both persons and dogs, more activation was seen in dorsal areas of the medial prefrontal cortex (mPFC), an area of the brain known to process emotion.

another study they conducted. They presented people with vignettes featuring a newly discovered alien species from another planet, called atlans. The “high experience” atlans were described as looking similar to slugs. They could feel pleasure and pain, but lacked intelligence, opinions, beliefs, and desires. The “high agency” atlans were described as looking similar to human beings except they were hard and metallic, like a humanoid robot. They could feel neither pain nor pleasure, but they did have thoughts, opinions, beliefs, and desires. Participants were then asked the same questions that were posed concerning the monkey experiment above: whether it would be wrong to capture the atlans for use in experiments, to use the atlans for human colonization, to kill the atlans, or to destroy “the atlan way of life.” There were in fact two variations of this study, one in which these questions were asked of the entire species, and one in which these questions were asked of a particular atlan. In both variations, responses to all four questions were significantly correlated with agential descriptions. Agential descriptions led people to show greater moral concern, regardless of whether participants were considering an individual atlan or the whole species. The experiential descriptions, by contrast, differed depending on the target. When considering the whole species, there was no significant effect. When considering an individual atlan, only the question of capturing the atlan was statistically significant (i.e., below the 0.05 threshold for statistical significance). The other three (using for human colonization, killing atlans, or destroying the atlan way of life) narrowly failed to reach statistical significance (i.e., were slightly above the 0.05 threshold and thus were not sufficiently related to showing those types of moral concern).

This study would appear to support Sytsma and Machery's hypothesis that Agency is also important for moral considerability. But there are problems here. First, the use of desires to indicate agency (without experience) is problematic. This would appear to be a nontraditional understanding of desire. Desires are typically understood to be an emotional or motivational state—something indicating valence. This is certainly the case in the ethical literature pertaining to animal desires (e.g., DeGrazia, 1996; Varner, 1998). Furthermore, in Gray et al.'s (2007) study desire is classified as an experiential state, possessing an Experience rating comparable with "Rage" and "Consciousness." Thus, in describing the atlans as possessing desires, Sytsma and Machery have introduced an element of Experience.

Second, Sytsma and Machery describe the high Experience atlans as unintelligent, nonsocial, and spending most of their time sitting on rocks, while the high Agency atlans were described as highly social and possessing "highly developed literary, musical, and artistic traditions, in addition to having made great advances in the sciences." These descriptions would appear to modify the nature of Experience in problematic ways. For instance, here being nonsocial is classified as an Experiential state, but no other study categorizes it in this way. Jack and Robbins, for one, describe experiential agents as being "emotionally responsive" and "warm towards others." Being completely nonsocial might indicate to people that they are unfeeling in a particularly morally salient way (even if they can feel other emotions, like pain and pleasure). The high Agency atlans, on the other hand, are described as social and interested in music and art, which implies significant Experience.

Of course, to clinch my argument against Agency, it would be nice to have supporting data, in addition to my criticisms of Sytsma and Machery's methodology. Nevertheless, given my criticisms, the best evidence suggests that agential states are unnecessary for basic judgments of moral considerability. Rather, the best explanation for the data presented thus far is that phenomenal states are both necessary and sufficient for attributing moral considerability to a being. When animals are attributed phenomenal mental states, they are also judged to be morally considerable. And when an animal is determined to be incapable of phenomenal states (or particularly unfeeling), no amount of Agency-related qualities lead people to judge the animal morally considerable. More research is of course needed to confirm these claims, but at present they are well-supported by the data.³³

However, someone might object that highly complex agential states are *sufficient* for moral considerability. As Sytsma and Machery (2012) say, "But Agency seems to matter too: People seem more willing to grant moral standing to animals whose cognition and lifestyle are complex. In fact, the more complex these are, the more people seem to be willing to grant moral standing."³⁴ A claim that seems to be widely shared

³³ Another study that might be taken to oppose the phenomenal account comes from Arico, Fiala, Goldberg, and Nichols (2011). 70% of participants in their experiment attributed anger, happiness, and pain to insects, but only about 6% did so to vehicles, clouds, or a blizzard, and 10% did so to plants. They did not ask any questions involving moral judgments, however, so it's not clear whether these acts of mentalizing would be sufficient to lead people to grant moral considerability to insects. This study is interesting, but at this point it is too inconclusive and anomalous to have any clear impact on the phenomenal account.

³⁴ Similarly, Martha Nussbaum (2006) has urged ethicists to adopt what she calls a "disjunctive approach" to moral considerability, which puts agential and experiential states on equal ground, "if a creature has *either* the capacity for pleasure and pain *or* the capacity for movement from

among ethicists is that agential states *entail* a certain degree of phenomenal states. This idea is illustrated by what I said above about music and art. These activities indicate the presence of complex cognition as well as phenomenal abilities. Something similar, we might think, is involved with other agential capacities, such as memory, language, and decision-making. Episodic memory, for instance, is typically thought of as being phenomenally rich (e.g., Varner, 2012). These states have an impact on an agent's moral importance because they entail a certain type of phenomenal state. Thus, in some cases high-level agential attributes are sufficient to lead to judgments of moral considerability.

I fully accept that certain (though not all) complex cognitive states can be sufficient for assigning moral considerability. This is consistent with the phenomenal account's necessity and sufficiency claims. In fact it supports the phenomenal account, because the reason certain complex cognitive states are sufficient for moral considerability is that they entail phenomenal states. This offers some reason for animal ethicists to emphasize agential states, but they would do so because of the phenomenal states entailed by *certain types* of agential states—namely, those of the right sort of complexity to entail changes in phenomenal experiences. Agential states that do not entail phenomenal states would not be sufficient to lead to judgments of moral considerability.

place to place *or* the capacity for emotion and affiliation *or* the capacity for reasoning, and so forth (we might add play, tool use and others), then that creature has moral standing” (p. 362).

3.2.3 The Role of Bodies in Phenomenal Mentalizing

A number of studies have found that an entity's body and behavior influence how it is mentalized. Here I will focus specifically on how phenomenal state mentalizing is influenced by such factors. The evidence, I suggest, indicates that a relatively narrow range of animals will be attributed phenomenal mental states. I will provide reasons to think this is problematic for assigning moral considerability.

That phenomenal mentalizing is suited to specific entities is evident simply from looking at which nonhuman entities scored highest in Experience on Grey et al.'s (2007) scale: apes and dogs. Both animals possess pronounced facial features similar to human beings, engage in complex social relationships, and display features (e.g. big eyes) that trigger the well-known "cute response" in human beings (Herzog, 2010). Apes can also occasionally walk bipedally and are phylogenetically close to human beings. All of these features tend to predict mentalizing of animals (Horowitz & Bekoff, 2007; Eddy, Povinelli, & Gallup, 1993).

Further evidence for the role of bodies in phenomenal mentalizing comes from Phillips and McCulloch (2005). They surveyed people from a large cross-national sample about their views on the sentience of various animal species. People were asked to rate the degree to which each species could feel pain, happiness, fear and boredom, in comparison to normal adult human beings. Their responses were then combined to create an aggregate sentience score. The final rankings, across all countries, went in the following order: monkey, dog, newborn baby (human), fox, pig, chicken, rat, fish. On the high end, monkeys, dogs, and newborn babies were judged to be about 80% as

sentient as normal adult human beings. At the bottom end, chickens and rats were judged about 60% and fish 47% as sentient as normal adult human beings. This study clearly indicates that sentience ratings are dictated by similarity to human beings.

Corroborating evidence that people's attitudes to animals are largely determined by similarity to human beings comes from Batt (2009). She presented people with pictures of 40 different species and asked them to rank how much they liked each species in comparison to all the others. She also created a ranking of each species according to their biological similarity to humans (which was calculated using a combination of behavioral, ecological, and anatomical information about each species). She then mapped people's preferences for each species onto the species' biological similarity to humans. The correlation between the two was quite high ($r = 0.542$, $P < 0.01$). Batt did not assess moral evaluations, but it is enough to show that biological similarity dictates general preferences. This indicates that people do indeed take such information into account.

As mentioned above, there is widespread agreement that all vertebrates are sentient, and thus morally considerable. But the evidence suggests people do not naturally attribute phenomenal states to all vertebrates. Rather, they attribute phenomenal states according to physical, functional, and phylogenetic similarity to humans.

To see the potential implications of this, consider the Phillips and McCulloch study again. This study did not ask any questions pertaining to morality, but one could argue that, if we had to set a moral considerability threshold according to sentience, the

75-80% range looks pretty accurate. Foxes and pigs were rated 65% and 67% respectively, but are arguably treated more like chickens and rats in contemporary Western society than monkeys and dogs (e.g., in being considered easily expendable). If this is right, and animals in the lower ranges are attributed “less” sentience because of their lack of similarity to humans, then a wide range of animals who are actually morally considerable (according to our attributions of phenomenal states) will fail to be evaluated accurately.

Even if this is not right, and the sentience-moral considerability threshold is much lower, the study nonetheless shows that similarity to human beings determines sentience ratings. This is enough for ethicists to be concerned. Suppose the threshold is at 50%, right above fish. Chickens, rats, foxes, and pigs would all be sufficiently sentient to be deemed morally considerable. Though this might seem significant, this fails to cover the thousands of amphibian and reptile species who physically resemble fish more than mammals, and would likely receive low sentience ratings.

One particularly problematic aspect of the way bodies and behavior influence mentalizing is that this process appears to be a type of ingroup bias, where we show a preference towards anything that possesses ingroup markers. This sort of psychological process is thought to be evolutionarily ancient and highly unmodifiable (Gil-White, 2001).³⁵ Human ingroups tend to use external physical features to identify group

³⁵ To be clear, what I am suggesting is unmodifiable in this paragraph is the *process* of using physical markers to distinguish ingroup and outgroup membership. By contrast, the actual physical markers used in the process of distinguishing ingroup and outgroup membership are modifiable at least in principle, though still resistant to change. So, for instance, the features used

membership (e.g., skin color, clothing, body modification). This applies to the case of animals as well. Animals' physical features are used to determine the extent to which they are part of our ingroup and thereby worthy of receiving phenomenal state attributions. A potential upshot of this is that human beings can learn new ingroup markers. So while we might currently use various mammalian or hominoid features to identify things "like us" (e.g., hairiness), perhaps in the future we could learn to use markers that apply to vertebrates more broadly. For instance, perhaps we could learn to use reptilian skin as a marker of sentience. While this might be possible, the evidence cited here suggests it is highly unlikely. The behavior and physical appearance of reptiles, amphibians, and many of the "lower mammals" are highly dissimilar from ours. The differences would appear to be much larger than that between most exclusive human groups. It would thus be incredibly difficult to convince people that animals bearing no resemblance to human beings should be deemed morally considerable.

The conclusion I take from this research is that animal ethicists who hope to convince people that a broad range of animals deserve moral consideration will meet great resistance. The mind assessment system we have inherited is primarily activated by organisms that are grossly similar to human beings. We are most familiar with mentalizing agents who are physically and functionally similar to ourselves, and so animals that bear little similarity to human beings will be denied the phenomenal mental

to distinguish ingroup membership could be changed with great investment of effort and resources, but no amount of resources can eradicate the process of basic ingroup/outgroup psychology.

states needed to be judged morally considerable.³⁶ In the next section I expand on the implications of the role of bodies in mentalizing, arguing that they impact moral significance judgments as well.

3.3 Phenomenal States and Moral Significance

Nearly every commentator on Gray et al. (2007) has interpreted their data as being relevant only to moral considerability (including Grey et al.). Nobody has yet asked whether the phenomenal account holds implications for moral significance.³⁷ The implications I see are relatively straightforward, stemming from the role of bodies in modifying phenomenal mentalizing as just outlined.

The basic idea here is simple. Animal bodies limit the range of entities assigned moral considerability, and an animal denied moral considerability has no moral significance. Regardless of how morally significant an animal's agential qualities *actually* are, they will only be relevant if the animal is first perceived to be capable of phenomenal experiences. The range of morally significant animals never expands beyond the range of morally considerable animals. And as I have argued, the role of animal bodies in phenomenal mentalizing narrows this range considerably.

³⁶ It has also been suggested that the robot in Sytsma and Machery's study is not ascribed phenomenal states because robots clearly lack the right type of bodily structures (Huebner, 2010a).

³⁷ As Sytsma and Machery (2012) point out, the categories of Experience and Agency map nicely onto dominant views in animal ethics. However, they apply these dimensions only to moral considerability. For instance, they explain, "An entity has moral standing if and only if it can be morally wronged. Thus, it is when, and only when, an entity has moral standing that the effects of a moral agent's actions on the entity directly...matter for the moral assessment of the actions. Entities that have moral standing deserve moral consideration, or concern, from moral agents" (p. 2) This is a textbook definition of moral considerability.

The animals that would seem to be most impacted by this feature of our psychologies are highly intelligent animals who do not physically resemble human beings. Dolphins, elephants, and various bird species are prime examples here. Though these intelligent species also express various human-like behaviors (e.g., sociality), they are very physically dissimilar from us. If they are not first attributed phenomenal mental states, their various complex cognitive capacities will never be brought to bear on their moral significance.

This point about the role of bodies follows directly from the above discussion of moral considerability, and I assume does not require further argumentation. What is more contentious is the role of complex cognitive capacities in determining the moral significance of animals *already* deemed morally considerable. Evidence is lacking here, but in the next subsection I will provide reasons to think that purely agential states will fail to have an impact on moral significance judgments, just as they fail with respect to considerability judgments.

3.3.1 Agential Capacities and Moral Significance

I suggested above that complex agential capacities are sufficient for moral considerability if they entail phenomenal states. They are not sufficient for moral considerability, however, if they are understood as purely agential. But what about for moral significance? Can purely agential capacities be brought to bear on the moral importance of animals already deemed capable of phenomenal experiences?

One high profile attempt to use cognitive abilities to argue for greater moral

significance is the great ape project (Cavalieri & Singer, 1993). Many who argue for the moral significance of great apes highlight apes' linguistic abilities (Savage-Rumbaugh, Shanker, & Taylor, 2001). These linguistic abilities, however, are usually described in purely agential terms. Kanzi the bonobo, for instance, is claimed to have an enhanced ability to communicate and think abstractly, but rarely is it claimed that his phenomenal experiences have changed as a result (though see Segerdahl, Fields, & Savage-Rumbaugh, 2005). The increase in moral importance demanded by apes' linguistic abilities is not claimed to be a result of any change in phenomenal states.

Though there might be reasons to support this idea, I am skeptical that ordinary human agents think language—understood as a purely agential capacity—is relevant to moral significance (nor for any other cognitive capacity). Unfortunately, there is not much evidence to appeal to on this question. Potentially suggestive results can be found in Knobe and Prinz's (2008) study, discussed above. For instance, only 9 percent of participants reported that studying a fish's memory would be related to moral concerns. The question posed to participants concerned morality in general, not just minimum considerability. This would seem to provide *prima facie* evidence against agential states being connected even to judgments of moral significance. However, there are problems. One is that people may not think fish are morally considerable, and so would not identify a memory question as pertinent to either considerability or significance. Another is that people might not think fish have good memories, and therefore would not grant them moral significance, even if they did think fish were morally considerable.

The other studies discussed above in support of the phenomenal account also

provide reasons to be at least somewhat skeptical of the role of complex cognitive capacities in determining moral significance. Nearly every study cited to demonstrate the ineffectiveness of agential states in determining moral considerability describes agents in terms of complex cognitive capacities. Planning and communication were highly rated as agential states in Gray et al. (2007); Bastian et al. (2011) described agents as “rational or logical”; and Sytsma and Machery’s (2012) highly agential monkeys were described as intelligent and inquisitive. In all cases, these descriptions failed to produce moral concern. The limitation in these studies is that they did not focus on animals that were already deemed morally considerable. This is what is needed to provide a proper assessment of the role of agential abilities in moral significance judgments.

Nonetheless, a tentative conclusion can be made: agential states do not determine moral significance judgments, absent inclusion of phenomenal states. More evidence is needed to substantiate this claim, however. I will discuss the implications of this further in the next section, where I outline the ethical implications of the phenomenal account.

3.4 Ethical Implications of the Phenomenal Account

In this section, I situate the preceding discussion in terms of psychological plausibility, as outlined in the last section. As stated previously, the clearest implication of the phenomenal account is that it is incongruent with pure agential accounts of moral considerability. I also stated that pure phenomenal accounts were supported in a somewhat trivial way, and that implications for mixed views would vary case-by-case. I will address all of these different types of theories here.

3.4.1 Psychologically Implausible Theories

The criteria of psychological plausibility most pertinent here are Impact, Achievability, and Transition. These criteria require ethical theories to have an impact on currently existing psychological profiles, or if they do not have any impact, their proponents must provide an account to explain how the theories might have an impact at some point in the future. In essence, we should ask whether theories in animal ethics can accommodate the evidence that currently existing human beings attribute moral considerability based on the ability to experience phenomenal states, or, if they cannot, whether they can provide a strategy for modifying the role of phenomenal states in human moral psychology.

Two main psychological processes I have outlined raise challenges for animal ethics. One is the role of animal bodies and behavior in constraining our phenomenal mentalizing, and the other is the absence of moral evaluations from agential mentalizing. Both of these processes exclude animals from moral considerability in cases where ethicists say otherwise.

These processes are particularly challenging because they are evolutionarily ancient and deeply engrained in our psychologies, and are thus extremely difficult to modify. Without significant intervention (e.g. a large segment of the population receiving years of experience with certain types of animals), the necessary changes would likely be unobtainable. First, we appear to be stuck with our bias towards things that look and act like us. This is problematic because the range of legitimately sentient animals is much broader than we are disposed to recognize. So while phenomenally-

based theories are, in general, psychologically plausible, they will be implausible to the extent they require people to assign moral considerability to a wide range of animals (as is required for most theories in animal ethics; Garner, 2013). This is not achievable for ordinary human beings.

One possible strategy for avoiding this limitation is to teach people about different types of animals, and show how the possession of phenomenal states is dissociable from whether an animal looks or acts like human beings. Huebner (2010b), for instance, objects to research on phenomenal mentalizing on the grounds that it does not allow for learning about animal bodies over time. He suggests that our perceptions of animal mental states will change as we come to know more about them.

Though being informed about animals certainly changes our mentalizing of them, I find it unlikely that this would be sufficient to overturn the phenomenal account. This education would need to have significant breadth and depth (enough to demonstrate that vertebrates are sentient) and would need to reach a wide audience (not just highly trained scientists). While this may be possible, it would take significant resources to successfully avoid our bias towards things that look and act like us. Overall, my assessment of theories that require people generally to attribute sentience to a wide range of animals is, firstly, that some theories will meet the minimum criteria for psychological plausibility, but will still need to overcome significant psychological limitations. And secondly, many other theories will not be minimally plausible without a rigorous transition strategy.

The most consistent finding of the research reviewed in this section is that

agential views are especially problematic. Agential states are not used to assign moral considerability by ordinary human beings. And while *some* people are clearly capable of using agential states to assign moral considerability (indeed, many philosophers!), clearly the intent of agential views (e.g., that of the early Frey) is for people generally to use agential states in this way. Given the evidence I have reviewed, that does not seem to be the case. Agential views thus appear to violate both the impact and achievability criteria.

It is also not clear how a transition to an agent-based moral psychology would look. Agent-detection is quite old, evolutionarily, and is also one of the first abilities developed in infancy (Arico et al., 2011). Physical cues for detecting agency include eye gaze, goal directedness, and contingent reactivity (e.g., responding flexibly to another agent). These are features even young infants (perhaps by 3 months) use to distinguish between agents and non-agents (Biro & Leslie, 2007). Attempts to modify moral considerability judgments to include attributions of agency would likely be building “on top of” these more fundamental agency-detection processes. Judgments of moral considerability based on phenomenal states, by contrast, seem to be “built into” phenomenal state attribution processes (according to the evidence reviewed above). Any proposal for transitioning to an agent-based moral psychology would thus need to explain how attributions of agency might come to be more strongly integrated with judgments of moral considerability, and also how attributions of phenomenal states that would otherwise drive these judgments might be suppressed. Overall, my assessment is that, given current evidence, agential views do not even meet the minimal criteria for

psychological plausibility.

Mixed views are less straightforward to classify in terms of psychological plausibility. There are two general types of mixed theories that would seem to be in question. The first includes agential states that entail phenomenal states (e.g., complex cognitive capacities like episodic memory and musical abilities). The second includes agential states applied to animals already deemed morally considerable (e.g., language in apes). Above, I suggested that agential states of the first sort can be sufficient for moral considerability judgments, and speculated that agential states of the second sort do not factor into moral significance judgments.

The crucial issue in both cases is achievability. With respect to entailment theories, we must ask *which* agential states entail phenomenality, and whether people's views on entailment can be changed. For example, consider Knobe and Prinz's experiment, in which a researcher was studying fish memory. If participants were interpreting this in terms of episodic memory—a phenomenal state—then, according to the phenomenal account, they would have connected the memory study to moral concerns. They did not, however, suggesting that they understood memory agentially. So it is possible that ethicists and ordinary agents will disagree about which agential states entail phenomenality. But what if it was made explicit that phenomenal states related to memory were being studied? Or perhaps a comparison was made between agential and phenomenal memory? Answers to these questions would illustrate whether people's concepts of agency are flexible with respect to entailing phenomenality. Despite the lack of data on this question, there are nonetheless good grounds to claim that entailment

theories pass the minimum standards for psychological plausibility.

Another type of mixed view that might be problematic is one in which agential and phenomenal states are combined but no phenomenal entailment is specified. Tom Regan's "subject of a life" criterion for moral considerability appears to fit this classification. His classic definition of being a "subject of a life" states:

Individuals are subjects of a life if they have beliefs and desires; perception, memory and a sense of the future, including their own future; an emotional life together with feelings of pleasure and pain; preference and welfare interests; a psychological unity over time and an individual welfare in the sense that their experiential life fares well or ill for them. (1983, p. 244)

This definition blends together a dizzying array of mental states, some agential and some phenomenal. He also says that subjects of a life "are individuals *who have an experiential* welfare – whose experiential life fares well or ill, depending on what happens to, or is done to or for, them" (Regan, 1983, p. 262). This makes his theory psychologically plausible, but it's not clear whether the other agential states he emphasizes would have a counteractive effect.

Things are similarly complicated with respect to agential views that do not entail phenomenality. The question here is whether ordinary human agents will use, or learn to use, pure agential states to assign moral significance. Answering this is difficult because the animal in question has been judged phenomenally capable, and thus any agential states might be *assumed* to have an impact on phenomenal states. A crucial test for this would be asking people to rank the moral significance of different phenomenally capable animals (or animal species) according to purely agential states. For instance, the Border collie Chaser has demonstrated the ability to comprehend over 1,000 English words

(Pilley & Reid, 2011). This would seem to be a purely agential ability. This sort of view about agential states could be tested by comparing Chaser's moral significance to an array of other dogs, all of which must be described as phenomenally capable but non-linguistic. My prediction would be that Chaser would not be judged to have greater moral significance, just as Kanzi the bonobo would not be judged to have greater moral significance than other bonobos. If, however, Chaser (or Kanzi) is granted increased moral significance, there would need to be an additional assessment of whether phenomenal states factored into their judgments. This would be relatively simple, and could be achieved by asking participants to fill out a basic mentalizing survey on Chaser and the other dogs. Without such evidence, however, it is difficult to assess the psychological plausibility of this view. It would seem to violate the minimum criteria, but the issue here is less that of needing a transition strategy and more needing data to assess how people actually use agential states to determine moral significance.

3.4.2 Concluding Thoughts

One upshot of this section is that it suggests animal ethics has been moving in a positive direction for the last 40 years. Phenomenal states have risen to particular prominence relatively recently, perhaps sometime in the 1970s in conjunction with work by Peter Singer and others on the concept of sentience (though of course there are many historical precursors, further substantiating the validity of the phenomenal account). The phenomenal account suggests that theories in animal ethics have thereby become increasingly psychologically plausible.

The issue with arguably the greatest significance here is the role of animal bodies in mentalizing. There are many questions unanswered concerning our ability to learn more about animals, and potentially learn to mentalize more accurately. This is particularly important because of what is at stake. If all vertebrates truly are sentient, then we are making a grave mistake in failing to attribute phenomenal states to them, thereby denying them moral considerability. So long as a creature is morally considerable, it is assured that its interests will be taken into account. This is absolutely fundamental, morally speaking. Future research on mentalizing would thus do well to achieve greater clarity on what makes us deny phenomenal states to legitimately morally considerable animals.

4. ANIMALS AND DISGUST

Among disgust researchers, the disgust-eliciting properties of animals are well known. The most influential research on the role of animals in causing disgust comes from Graham Davey and Paul Rozin, both of whom place animals at the center of the evolved functions of disgust. Rozin's theory of disgust, which is arguably the most widely accepted over the last 25 years, places special emphasis on animals. As he and his colleagues claim in a recent survey article, the evolved function of disgust is to protect us from seeing ourselves as "lowered, debased, and mortal" (Rozin, Haidt, & McCauley, 2008, p. 762). This includes being in the company of animals, because "anything that reminds us that we are animals elicits disgust" (p. 761).

Despite the centrality of animals in the history of disgust research, the most prominent philosophical accounts place animals in the periphery. McGinn (2011), for instance, makes little mention of the disgust-eliciting properties of animals, focusing only on the symbolic nature of animal disgust (that we do not wish to see ourselves as animal-like). Dan Kelly's (2011) influential account, though it is strongly rooted in evolutionary history and is broadly Rozinian in its claims, does not investigate the role of animals in any detail.

This section attempts to fill this gap in the literature. I draw from Rozin's theory of disgust, as well as Kelly (2011), to argue that animals elicit disgust in two ways. One is by triggering disease-protection mechanisms, and the other is by eliciting what is called mortality salience, or thoughts of death (as indicated in Rozin's comments above). Both of these indicate that animals cause aversive and avoidant reactions in human

beings.

These aversive reactions caused by animals raise two further challenges, one pertaining to moral psychology and another pertaining to ethics. With respect to the former, it seems odd that animals would cause disgust, given their prominence in human lives. I attempt to account for this by utilizing research by Nick Haslam (2006) on infracommunication and dehumanization. I argue that this research suggests that we cope with animals, despite their disgust evoking powers, by attributing them mental states that include positive evaluations but simultaneously function to cement animals' status as inferior beings. While I think this adequately addresses the psychological challenge, it indicates that improving treatment for animals will be incredibly difficult. Both the aversiveness of animals and the way we "dehumanize" them present significant psychological obstacles for ethicists aiming to improve the moral status of animals. I conclude by discussing recent objections to the importance of disgust in normative ethics raised by those Dan Kelly calls "disgust skeptics." I draw from the discussion of psychological plausibility in section 2 to respond to these objections.

4.1 Animals as Disgust Elicitors

4.1.1 Core Disgust

I will begin by outlining animals' role in what is known as core disgust. Core disgust is so named because it is hypothesized to represent the evolutionary functional core of disgust. This is a natural starting point for providing an account of disgust, given that animals have a constant presence in human evolutionary history.

Rozin et al. (2008) argue that core disgust consists of three different features: oral incorporation, a sense of offensiveness, and contamination. As Kelly (2011), following Rozin, characterizes the emotion, disgust consists of “aversion with a pronounced oral feel” (p. 17). The evolutionary function of disgust, on this account, is to protect against potential contaminants.³⁸ The primary way of doing this is by blocking contaminants from invading the body, particularly through the mouth. This is what is meant by oral incorporation.

It might be obvious why animals are relevant here. As food and as vectors of disease, animals have served as a primary cause of death and illness in our evolutionary history. This is why core disgust functions to keep us away from certain animals, and particularly to avoid putting some of them in our mouths.

That we would develop a disgust response to animals makes sense when thinking about the history of disease transmission. For instance, of the 25 most deadly diseases in human history, 9 are due primarily to animal vectors, and in 20 of the 25 cases the transmission vector is classified as core disgust—things like feces, bodily fluids, lice, flies, and other bugs (Oaten, Stevensen, & Case, 2011; also see Curtis and Biran, 2001). This provides a reason to think that animals factor centrally in disgust. It’s also interesting that all 25 of these diseases possess behavioral indicators of the disease in those infected, 23 of which can be observed from looking at the face. So not only are animals the cause of various diseases, and so likely to be at the core of disgust, but the

³⁸ Davey and Rozin proposed early on that disgust functions as a disease-avoidance mechanism (Davey, 1992, Rozin & Fallon, 1987)

diseases they cause are also highly public. This makes it likely that people will be aware of the link between animals and the diseases they transmit to human beings.

Other studies have shed light on disgust-eliciting food, particularly with respect to meat. Martins and Pliner (2006) found that unfamiliar foods are rated as disgusting largely to the extent that they have animal origins. This would seem to be a result of a disgust-disease-avoidance mechanism. Other suggestive evidence comes from Fessler and Navarrete (2003), who looked at food consumption in 78 different cultures and found that meat was the most highly regulated food item. Given the possibility of infection from eating meat, it would make sense for people to regulate its consumption, and for disgust mechanisms to be driving such regulation.

Given this evidence, it would appear that animals factor into core disgust either by being consumed or by transmitting disease through some other entryway into the human body. This dual role of animals in causing disgust has also been pointed out by Kelly (2011). Kelly's account of disgust diverges slightly from Rozin, but the differences do not raise any significant issues that need addressing here. I will briefly describe Kelly's account as complementary to Rozin's in order to illustrate how this dual role might work.

On Kelly's account, disgust consists of two "entangled" mechanisms: the poison mechanism and the parasite mechanism. They were once independent, according to Kelly, but in human beings are currently combined as a single disgust response. The poison mechanism responds to poisonous substances we might put in our mouths, and is responsible for providing the affect of disgust—what we feel. The parasite mechanism

responds to things that indicate the presence of toxins—things like parasites and other pathogens. As Kelly points out, animals are important here because of their relevance to both categories. Animals are frequently ingested and frequently carry parasites and other pathogens. Simply the presence of animals might be a health risk, even if we are not in direct contact with them.

Kelly's account helps make sense of why we would avoid animals. As Kelly explains, the disgust system is geared to be overly sensitive and to produce a high rate of false positives. A single false negative—failing to identify a legitimate risk—could be fatal, so it is worth erring on the side of caution. Many diseases are airborne, and, as we will see below, many are carried by flying insects. These factors are likely to be represented in our disgust responses, leading us to exaggerate our avoidance of animals and steer clear of them more than is necessary. For instance, simply avoiding their ingestion, via the poison mechanism, would not be sufficient. This avoidance response, as well as other functions of disgust, will be explained in the next section.

4.1.2 How Disgust Works

As I just mentioned, one main function of disgust is to be overly sensitive to potential contaminants, in a way that produces many false positives. There are a number of well-known examples of this: chocolate can elicit disgust if made to look like feces, and people will refuse to drink out of a glass that previously contained a dried and sterilized cockroach, even if they are told that doing so would be safe (Rozin, Millman, & Nemeroff, 1986). This is just one of many features of disgust that suggest the disgust

response is automatic and invariant to a significant degree. Once the disgust mechanism is triggered, it cannot be modified. The system flings itself into action to protect against the contaminant, rather than work to determine whether the contaminant does in fact carry the risk of infection. As Kelly (2011) describes it, the disgust system is “elicitor neutral” (p. 19) and “unified” or “clustered” (p. 40) in its response. In essence, this means that nearly anything can be made disgusting, and once that status is achieved, a chain of responses follows automatically.

I will describe three different behavioral responses that are part of the disgust system. Each of these responses appears to be automatic and invariant, as just described. These do not exhaust the disgust response, but they are essential functions of the system, particularly in responding to animal triggers.

The first response is protection of the “bodily envelope.” Though many disgust researchers agree that disgust is the foremost protector of the mouth, it also functions to protect any gateway to the body. For instance, in a classic study, Rozin, Nemeroff, Horowitz, Gordon, and Voet (1995) asked people how they would feel if either a clean Q-tip or a Q-tip that had been licked by a stranger was placed in or on different parts of their body. The parts of the body that elicited the strongest negative emotions were those that serve as gateways to the body, primarily the mouth and genitalia. People are also more disgusted by the thought of transplanting “interface parts,” or those that serve the gateway function (like the mouth), than they are other internal parts of the body (Fessler & Haley, 2006).

The second response expands on the oral incorporation feature already described. Since disgust primarily protects the mouth, as the chief gatekeeper of the body, it may be no surprise that the disgust response activates the entire digestive system. Even things that are not normally consumed, such as blood and feces, still activate a disgust response that employs the digestive system. For instance, nausea and increased salivation are part of the disgust response, as is the classic “gape” face (think about the way someone’s face looks right before they begin to vomit). Anyone who has experienced extreme nausea or has spent extended periods of time in a state of near-vomit knows that these experiences are highly noxious and unpleasant. This is important to note because these responses are part of *any* disgust response, not just those related to ingestion. So, for example, someone who is disgusted by animals will undergo this sort of visceral and gustatory response, regardless of whether or not they consume animals or have consumption-specific disgust.

The third response follows from the two already mentioned. Perhaps the most crucial feature of disgust is that it is an aversive emotion. It is hypothesized to be driven by two different systems with the same acronym: the Behavioral Inhibition System (Carver & White, 1994; Olatunji, Haidt, McKay, & David, 2008) and the Behavioral Immune System (Schaller, 2006; Schaller & Duncan, 2007; Schaller & Murray, 2008). Both systems are important in that they motivate withdrawal. The Behavioral Immune System is involved with protection from potential contaminants of both a biological (e.g. bodily fluids) and social nature (e.g., people who one finds disgusting). The Behavioral Inhibition System includes a variety of emotions (like fear and sadness) that motivate

withdrawal and removing oneself from whatever is triggering the system. For disgust, the trigger is potential contagion. For instance, the response that follows from being disgusted by animals is dislike, aversion, and avoidance. As might be obvious, the reason it does this is to protect the bodily envelope, particularly the mouth. So not only does disgust produce high rates of false positives, it also produces exaggerated avoidance responses. This ensures adequate protection from disease.

A key question, of course, is *which* animals elicit disgust. It is not the case that all animals elicit the same level of disgust, and some animals possess features that can override our natural disgust response. These issues will be explored in the next section.

4.1.3 Which Animals Elicit Core Disgust?

Identifying the ethical implications of disgust responses requires us to more clearly outline which animals cause which types of avoidance responses. A taxonomy widely used by disgust researchers distinguishes between predators, who primarily elicit fear, slimy invertebrates, who primarily elicit disgust, and a category in between these two containing a wide variety of different animals that elicit both fear and disgust. Tigers and sharks, for example, elicit fear, while snails and maggots elicit disgust, and rats, cockroaches, and spiders elicit both fear and disgust (Matchett and Davey, 1991; Ware, Jain, Burgess, & Davey, 1994; Webb & Davey, 1992; for cross-cultural support see Davey et al., 1998).

What is important in providing an account of animals as disgust elicitors, however, is proving that they cause disgust because they are linked to disease and other

pathogens. Roughly, what we should look for is evidence that disgust responses to animals are related to disease avoidance, or confer some sort of health benefit. I will discuss three experiments that help illustrate such a relationship.

One widely cited study comes from Curtis, Aunger, and Rabie (2004). They conducted a large cross-cultural study (over 40,000 participants), comparing people's reactions to a variety of different disgusting photos. Some of the photos were disease-relevant while others were disease-irrelevant, but participants were not informed of this difference. For instance, lice, which can transmit diseases to humans, were compared to wasps, which cannot. Another example is *Ascaris* worms (disease-relevant), which were compared to caterpillars (disease-irrelevant). The results showed that people indeed rated the disease-relevant photos as more disgusting than the highly disgusting but disease irrelevant photos. This would seem to provide evidence in favor of the disgust-disease relationship.

Another study comes from Prokop, Fančovičová, and Fedor (2010). They presented people with the pictures from Curtis, Aunger, and Rabie's experiment (some of which were disease-relevant and others that were not), and asked them to rate the extent to which they were afraid or disgusted by each picture. Participants also took a standard health survey that included questions about the participants' antiparasite behaviors, some of which were relevant to animals. For instance, one question asked, "Do you pet wandering/home cats/dogs?" As expected, the disease-relevant insects elicited more fear and more disgust than the disease-irrelevant insects. Interestingly, people who were more disgusted by disease-relevant insects as well as those who were

more fearful of disease-irrelevant insects also reported more antiparasite behaviors. Though the animal-related questions did not receive their own analysis, these results suggest those who are disgusted by and afraid of insects are likely to engage in animal avoidance more broadly.

The third experiment addresses disgust oriented towards domestic pets. Prokop and Fančovičová (2011) asked pet owners to fill out a survey designed to assess their vulnerability to disease. Participants were also asked to rate four additional questions pertaining to their pets, two relevant to parasite transmission (“I allow my dog/cat to lick me,” “I allow my dog/cat to sleep in my bed”) and two associated with parasite avoidance (“I worm my dog/cat regularly,” “I frequently get rid of my dog/cats feces”). They found that pet owners generally perceive themselves as having low vulnerability to disease in general. Interestingly, however, parasite transmission was positively correlated with parasite avoidance. That is, pet owners who made themselves more available to parasite transmission through close contact with their pets also made greater efforts to ensure that their pets were free of parasites. This suggests that pet owners perceive themselves to be less sensitive to disgust elicitors, except when they are knowingly putting themselves in direct contact with potential pathogens, just as the core disgust account would predict.

4.1.4 Implications of Core Animal Disgust

Before moving onto the second type of animal-elicited disgust, I will briefly review the implications of the research I have presented thus far. I take the experiments

cited here to show 1) a wide range of animals reliably elicit disgust in human beings, and 2) they do so because of their relationship to disease transmission. The main problem pertaining to ethics arises out of the avoidance response produced by core animal disgust. People are likely to avoid animals and refrain from, say, providing assistance to them, insofar as doing so would require them to be in contact with animals. Proactively helping animals requires approach—not avoidance—motivations. If an animal is ill, disease-ridden, or injured, it is likely to possess even more of the disgust elicitors it would normally express. An amplified disgust response would be predicted on the part of human beings in such cases, which is likely to lead to continued neglect and denial of assistance. That is, our disgust mechanism predicts avoidance of animals precisely in the cases where animals are likely to require the greatest assistance.

This avoidance response, when combined with other evidence reviewed thus far, presents a significant challenge for ethicists. Our disgust response to animals is likely to be exaggerated and contain many false positives, it will be engaged simply by animals' presence, regardless of whether they are for consumption or might actually be placed near our mouths, and it is invariant and automatic. Once an animal is perceived to be disgusting, the avoidance response follows automatically. Disgust researchers also generally agree that disgust is highly unmodifiable and unresponsive to cognitive control (e.g., Russell & Giner-Sorolla, 2013). Core disgust elicitors are hard to unlearn, even with habitual exposure. For instance, treatment for spider phobias tends not to generalize beyond spiders (Smits, Telch, & Randall, 2002). In one suggestive study from Rozin (2008), the disgust responses of medical students who regularly dissected cadavers were

only reduced over time in response to cold bodies, but not recently deceased warm bodies. From this, Rozin concluded that learning new information about a disgust elicitor is uninformative about any other disgust elicitor. With respect to animals, this suggests that reducing one's disgust response for any one particular type of animal will fail to generalize to any others.

But, as is clear from these studies, the primary disgust-eliciting animals tend to be invertebrates and insects (particularly the slimy ones). For many animal ethicists, these animals do not possess moral status anyway (because they are not sentient) and so do not cause any problems for their theories. Spiders, for instance, are commonly cited as the most widely disliked creature in the animal kingdom (Davey, 1994; Davey & Marzillier, 2009), but spiders rarely feature heavily in discussions of moral status. Among vertebrates, the primary disgust eliciting animals are rodents, while farm animals, for instance, appear to cause more fear than disgust (Arrindell, Mulkins, Kok, and Vollenbroek, 1999). While rodents are certainly important for animal ethics, they arguably are not near as central or salient as farm animals and other "higher mammals." And these do not appear to cause a problematic level of disgust.

The obstacles caused by core disgust might thus appear to be relatively innocuous. The level of disgust elicited by the most morally significant animals will likely produce some level of avoidance, but not near as much as towards invertebrates and other organisms lower in moral significance. From the evidence cited thus far, core disgust does not obviously present insurmountable problems for animal ethics. However,

in the next section I turn to the second category of disgust elicited by animals, and as we will see, the obstacles there are much more significant.

4.2 Animal Reminder, Mortality Salience, and Terror Management

4.2.1 Animal Reminder Disgust

As mentioned already, the Rozinian account of disgust claims that disgust protects us from seeing ourselves as “lowered, debased, and mortal.” The discussion of core disgust above, however, characterized disgust as a protector against various pathogens. So what connection is there supposed to be between contamination and mortality?

The line of research that has pursued this connection in greatest detail is known as terror management theory (TMT). TMT primarily derives from work in anthropology by Ernest Becker (1971, 1973, 1975). Within psychology, TMT was first developed by Sheldon Solomon, Jeff Greenberg, Thomas Pyszczynski, and colleagues (Greenberg, Pyszczynski, & Solomon, 1986; Greenberg, Pyszczynski, Solomon, Simon, & Breus, 1994; Rosenblatt, Greenberg, Solomon, Pyszczynski, & Lyon, 1989; Solomon, Greenberg, & Pyszczynski, 1991). TMT starts from the assumption that human beings are regularly confronted with and experience acute feelings of existential anxiety. It is important for normal functioning that these feelings stay beneath the surface of consciousness, so we have developed various methods for repelling these feelings when they bubble to the top (for reviews of TMT see Burke, Martens, & Faucher, 2010; Goldenberg, Pyszczynski, Greenberg, & Solomon, 2000; Greenberg, Solomon, &

Pyszczynski, 1997; Hayes, Schimel, Arndt, & Faucher, 2010; and Solomon, Greenberg, & Pyszczynski, 2004).

The hypothesized reason we possess this anxiety, at least with respect to animals, is that animals remind us we are mortal. This is what is referred to as animal reminder or mortality salience. According to Rozin and his colleagues, disgust is both “a defensive emotion that guards against the recognition of our animality” (Haidt, McCauley, & Rozin, 1994, p. 712) and “a defense against a universal fear of death” (Rozin, Haidt, & McCauley, 2000, p. 643). The result of these supposed threats is that human beings “wish to avoid any ambiguity about their status by accentuating the human-animal boundary” (Rozin and Fallon, 1987, p. 28). This process is perhaps best explained by Goldenberg, Pyszczynski, Greenberg, and Solomon (2000):

The terror management solution to the problem of death is to live our lives on an abstract symbolic plane: We cope with the threat of death by embedding ourselves in a meaningful culture and living up to the culture’s standards. In this way, we elevate ourselves above the rest of the animal kingdom. (p. 203)

The basic connection between disgust and TMT is that disgust helps us avoid those things that cause us existential anxiety: animals.

One way of testing TMT is by manipulating people’s thoughts of death (thus modifying mortality salience). Making death more salient increases people’s existential anxiety, or so the hypothesis runs. The traditional method of doing this is by presenting participants with a death prompt. The most commonly used death prompt asks participants to describe the emotions they feel when they think of their own death and what they think will happen, physically, when they die. Other popular and more implicit death primes include reading articles about death, word searches and sentence

completion tasks that include death-related words, and interviewing participants in or near a cemetery or funeral home. The typical result of the death prompt is that people are inclined to rate death and disgust-related items negatively.

Though these death primes might seem somewhat artificial and uncommon in everyday life, TMT researchers generally suppose that our lives are filled with more mundane death primes that function in essentially the same way. For instance, other effective elicitors of mortality salience include cancer (Arndt, Cook, Goldenberg, and Cox (2007), breast exams (Goldenberg, Arndt, Hart, & Routledge, 2008), 9/11 (Landau et al., 2004), old people (Martens, Greenberg, Schimel, & Landau, 2004), people with disabilities (Hirschberger, Florian, & Mikulincer, 2005), and fecal matter (Dunkel, 2009), among many others. These are relatively common. The prevalence of these mundane death primes makes it especially problematic that animals too function to prime death-related thoughts.

What comes after the death prompt in TMT studies generally varies according to other variables being measured. Rather than test anxiety as such, TMT researchers typically measure the predicted response to repelling this anxiety (though direct studies of anxiety and other mental states exist as well; e.g., Routledge, Ostafin, Juhl, Sedikides, Cathey, & Liao, 2010). The most well-confirmed methods for repelling the anxiety produced by mortality salience include 1) affirming one's core values and beliefs, or the values and beliefs of one's culture, and 2) increasing one's self-esteem (Gailliot, Stillman, Schmeichel, Maner, & Plant, 2008; Harmon-Jones et al., 1997; Schmeichel et al., 2009; Schmeichel & Martens, 2005).

I will briefly discuss three experiments illustrating that animals cause death-related thoughts and that this process incorporates disgust. In a classic study, Goldenberg et al. (2001) presented participants with the traditional death prompt, followed by a disgust assessment. As the core disgust account would predict, the death prompt led to higher ratings of disgust for body products as well as maggots, worms, rats, and cockroaches, but not other types of disgust elicitors. This indicates that mortality salience does indeed activate disgust mechanisms specific to animals.

Another illustrative experiment comes from Cox, Goldenberg, Pyszczynski, and Weise (2007). They found that presenting people with descriptions of core disgust items (e.g., “It would bother me to see a bowel movement left unflushed in a public toilet”), when combined with a human-animal similarity prime, led to increased use of death-related words in a word completion task. The core disgust prime was insufficient on its own to produce increased thoughts of death, indicating that the human-animal similarity prime played a crucial causal role.

Perhaps the best evidence in support of the claim that animals elicit mortality salience comes from Beatson and Halloran (2007). They presented participants with the traditional death prompt as well as an animal stimulus, which was a video of bonobos having sex. In the High Creatureliness condition, as they called it, the researchers emphasized to participants how similar bonobo sex is to human sex. In the Low Creatureliness condition, the *differences* between the species’ sex habits were emphasized. Participants’ self-esteem was measured, as were their attitudes to animals. The results were consistent with TMT and indicate that bonobo sex elicits mortality

salience. Among those in the High Creatureliness condition, where bonobo sex was compared to human sex, those with low self-esteem evaluated animals more negatively, while those high in self-esteem evaluated animals more positively. These results are exactly as would be predicted by other classic TMT studies.³⁹ Reminding people of their animal nature caused negative evaluations, unless the level of self-esteem was sufficient to fend off the attendant existential anxiety.

These experiments indicate that animals do indeed function as mortality salience enhancers, and that this leads people to judge animals negatively. I will outline the implications of this research for animal ethics in a later section. First, however, I will address some criticisms of TMT, particularly with respect to animals.

4.2.2 Criticisms of Terror Management Theory

There have been a wide variety of criticisms of TMT, many of them aimed at the role of animals. While the results themselves seem extraordinarily robust, the theoretical foundations of TMT are quite shaky. A well-known article by Tybur, Griskevicius, and Lieberman (2009) identifies a number of the pertinent theoretical problems. One is that it's not clear why existential anxiety would have any adaptive function, particularly anxiety caused by animal reminders. It is incapacitating, and it's not clear why or what it helps us accomplish. It's also not clear why a disgust response, appealing to one's worldview, or increasing one's self-esteem would be particularly suited for repelling existential anxiety. Lastly, it's not evident that people actually avoid being reminded of

³⁹ For a replication with pet-owning participants see Beatson, Loughnan, & Halloran (2009).

their animality. In fact it seems unlikely that animals would be a source of anxiety, given the close proximity of animals and human throughout our history (see Navarrete, 2005 and Kirkpatrick & Navarrete, 2006 for similar criticisms of TMT).

For the most part, I find these criticisms compelling. However, I will address one additional criticism made by Tybur et al. stemming from those just mentioned. Tybur et al. argue that there isn't any feature in common between different types of animal reminder stimuli. Sex, bad hygiene, corpses, and flesh wounds all elicit animal reminder disgust and are classified as such, but these factors do not seem to have anything in common with animals (or each other). If anything, they are all pathogen related, not reminders of our animality. Furthermore, avoidance of pathogens has a clear evolutionary purpose, and being sensitive to animal reminders does not.

My response to Tybur et al. is that the unifying feature of animal reminder elicitors is simply that they all cause a certain expected response. That is, they all cause anxiety that is alleviated by broadening the human-animal boundary, by affirming one's worldview, and increasing one's self-esteem. This is a quite specific response that is not caused by a number of other things, including a number of other disgust elicitors. For instance, reading about animal *abuse* does not increase thoughts of death (Das, Bushman, Bezemer, Kerkhof, & Vermeulen, 2009).⁴⁰

⁴⁰ Or, consider another classic finding in TMT research, that those who score high in neuroticism (according to traditional clinical diagnoses) tend to rate the physical aspects of sex much more negatively compared to the romantic aspects (Goldenberg, Pyszczynski, McCoy, Greenberg, & Solomon, 1999). Interestingly, this can be replicated in non-neurotics simply by telling them that humans are similar to animals. It can also be reversed in neurotics by telling them that humans

It may not be obvious *why* they would share this feature, but this is not necessarily detrimental to the theory. Though I share Tybur et al.'s skepticism of TMT's theoretical foundations, the evidence in its favor is hard to deny. There are hundreds of other studies, besides the ones I have already reviewed, confirming the basic role of mortality salience in causing anxiety that is alleviated primarily by increasing one's self esteem and affirming one's worldview. Many of these studies also test a wide variety of alternative explanations, none of which have proved tenable (e.g., increased arousal or other sources of anxiety; see Goldenberg et al., 2000 and Greenberg et al., 2008 for a review). In short, we can accept the basic phenomena of TMT without accepting its full theoretical implications.⁴¹

A remaining issue with TMT as it relates to animals is precisely how problematic animals specifically—and not other animal reminders, like dead bodies—are for our evaluations of animals. Dead and diseased animals elicit both core and animal reminder disgust, but this is relatively unsurprising. There is also some evidence that animal reminder disgust fails to be elicited by animals as much as would be expected. For instance, Fessler and Navarrete (2005) report that Costa Ricans are exposed to dead and decaying animals more than people from most countries, but do not appear to have heightened animal reminder disgust. And in one study (Burris and Rempel, 2004),

are *different* from animals (Goldenberg, Cox, Pyszczynski, Greenberg, & Solomon, 2002). This would seem to provide good evidence for TMT (and the role of animals within the theory).

⁴¹ Various other evidence supports the claim that animal reminder disgust functions as hypothesized within TMT. For instance, it's notable that animal reminder disgust has consistently been found to be a dissociable component of the disgust response, even cross-culturally (Olatunji, Haidt, McKay, & David, 2008; Olatunji et al., 2009).

people who read a graphic description about dust mites subsequently increased their worldview defense but did not show an increase in thoughts about death. So how much, exactly, do animals cause us existential anxiety in a way that leads us to treat and evaluate them negatively? Can the effects of animal reminder disgust be strongly separated from core disgust?

Notwithstanding the evidence I have described, in many cases I suspect we will not be able to tell whether disgust towards animals is due to their mortality salience or their status as potential transmitters of disease. Fortunately, this is not problematic. It is enough to know, as described previously, that animals elicit aversive behaviors in human beings. The discussion of core disgust suggested that this might be limited to the “lower” animals, such as invertebrates, but the research presented here on animal reminder disgust should illustrate that the disgust response towards animals will encompass many “higher” animals as well. Studies on TMT indicate that animals—even pets—can elicit negative evaluations. We do not have to accept that the reason for this is that they cause us existential anxiety. It is enough to focus on the effects of animals’ presence. They appear to increase thoughts of death, which, for whatever reason, lead us to judge animals negatively.⁴²

⁴² Someone might object at this point that I have focused too much on biological disgust and too little on moral disgust. However, there is a rich empirical literature on the connection between the two (e.g., Chapman & Anderson, 2013; Eskine, Kacinić, & Prinz, 2011; Schnall et al., 2008; Tybur et al., 2013). To briefly summarize, the evidence suggests that the two are part of the same system, and that biological disgust in many ways drives our moral evaluations. Current thought on why moral disgust would be connected to biological disgust is that they *both* function to protect against contaminants (Rozin, Haidt, & Fincher, 2009). Activation of either system elicits the same facial expressions (Chapman, Kim, Susskind, & Anderson, 2009), the bitter taste rejection response (Rozin, Haidt, & Fincher, 2009), and the same avoidance action tendencies

4.3 Dehumanization and Infracumanization

The widespread presence of animals in human lives seems somewhat paradoxical, given the research discussed thus far. Animals tend to trigger feelings of disgust and thoughts of death in human beings, leading us to avoid them and judge them negatively. So how do we explain their presence in our lives? How do we cope with them, given that they elicit disgust? I will address these questions in this section in the context of research on dehumanization and infracumanization.

Some of the most exciting research on our moral evaluations of nonhumans has come from studies of how we use the classification of “animal” to demean other human beings. Various studies have shown that this process is primarily driven by feelings of disgust and frequently targets outgroups. This maps well onto research on perception of nonhumans. To consider some highlights: Boccato, Capozza, Falvo, and Durante (2008) found that subliminal presentation of monkey faces facilitated identification of outgroup names but human faces did not; Costello and Hodson (2010) found that people who perceived a greater divide between humans and other animals were more likely to engage in racial dehumanization (among human beings); and Buckels and Trapnell (2013) found that experimentally induced disgust produced stronger implicit associations between animals and an outgroup than did induced sadness or induced neutral mood.

(Carver & White, 1994). Biological and moral disgust also activate a similar network of brain regions (e.g., the globus pallidus, putamen, caudate head, and amygdala; Schachar, Borg, Leiberman, & Kiehl, 2008).

In general, there are two types of processes people use in these sorts of classifications (Haslam & Loughnan, 2014). One is called dehumanization, which refers to the process of explicitly identifying other people (or groups of people) as nonhuman. The other process is called infrahumanization, which refers to the process of identifying other people (or groups of people) as *inferior* human beings. While they are still attributed various key human qualities, they are treated as inferior to some other group by comparison. What is surprising in infrahumanization is that it produces a certain level of positive evaluation. It is not thoroughly intended to demean others, just enough to make it clear who is superior. This maps well onto how we treat some nonhumans. We live with them and show them care, but in many cases, continue to see ourselves as dramatically superior.

In this section, I will make a proposal. The discussion of disgust thus far has raised a problem for explaining the role of animals in human lives. Namely, it's not clear why we ever show them moral concern or affection, given that they are universal disgust elicitors (along multiple dimensions). The processes of dehumanization and infrahumanization potentially illustrate why this is and how this works. Animals, just like human outgroups, cannot really be avoided. We must live with them. Animals and humans were in close contact long before domestication (Serpell, 1996). And it's hard to imagine how anyone in contemporary societies could avoid animals entirely. People are thus forced to live among disgust elicitors.

The prediction would be that some sort of coping mechanism must be in place for this. A strategy that comes out of the TMT discussion above is that we would find

ways to elevate ourselves above animals. This in itself is not sufficient, however, since animals are pervasive in contemporary society, so we would need to apply this coping mechanism constantly. It also fails to explain why we frequently find animals appealing and judge them positively. Rather, I think we should look to the positive evaluations that result from dehumanization and infrahumanization for assistance. On the face of it, this too is a puzzling feature of our disgust response. The role played by disgust in these processes is to remove “contaminated” others from ourselves—consistent with the biological origins of disgust. Yet dehumanized others still receive various positive attributions. The solution I see is as follows. We can’t always avoid those we have dehumanized or find disgusting, such as animals. For instance, in human evolutionary history it would have been costly to avoid animals, regardless of the disgust they caused. Attributing various positive human qualities to dehumanized others—sometimes even to the point of explicit exaggeration (as in anthropomorphism)—makes it easier to live among things that elicit disgust.

Developing this proposal requires further discussion of dehumanization and infrahumanization. Leyens et al. (2001) was the first to propose roughly the classification system described above. Haslam (2006) refined this further to identify two further processes that lie at the heart of dehumanization and infrahumanization. One type, called *human uniqueness*, denies to another entity properties that are considered uniquely human. This is what is often denied to animals. Accusing others of being shameless, lacking humility, or being unsophisticated are common examples. The second type, called *human nature*, pertains to properties that are essentially human,

though they may not be unique to humans. Common examples include accusing someone of being insensitive, unreflective, or reckless. These tend to be denied to robots and machines.

So what this predicts is that animals, and various human outgroups, will be denied uniquely human traits but still be attributed essential human traits. What actually happens, however, is that animals in particular are attributed these essential traits to a *greater degree* than other human beings. In one widely cited study of this (Haslam, Kashima, Loughnan, Shi, & Suitner, 2008),⁴³ animals were judged to be more perceptually capable than humans, in having a greater ability to hear, smell, taste, and see. Their ability to experience basic emotions (like anger, disgust, excitement, fear, and surprise) and volitional states (like needing, wanting, and willing) were judged to be nearly equivalent to human beings. The states that were denied to them were higher cognitive states (like knowing, reasoning, and imagining), and more complex emotions (like admiration, enjoyment, love, pride, and resentment). So animals are not as reflective as human beings and lack complex emotional abilities, but are more perceptual and possess similar basic emotions. It is problematic, of course, that Haslam et al. did not specify to participants which animals were in question, but this basic phenomenon is interesting nonetheless.

One prominent feature commonly cited in explaining why we like animals and (for some of us) bring them into our homes is that they provide companionship. Indeed,

⁴³ The sample consisted of participants from Australia, China, and Italy. 37 mental states total were measured for animals, robots, and supernatural beings.

recent studies by Epley, Waytz, and Cacioppo (2007) show that people who score high in a need for social connection are more likely to attribute mental states to animals. The problem I have highlighted is that the disgust eliciting powers of animals make this seem strange. But by hypothesis, even those who like animals for companionship would be bothered if their pets were not clearly relegated to an inferior class of beings.

Given the pervasiveness of pets, it is helpful for us to have a stock of anthropomorphic terms to describe their behaviors. The proposal I am making is that this is precisely what allows us to not feel threatened by animals. The particular mental states we attribute them solidifies their infrahumanized status. They are evaluated positively (e.g., being adept perceptually) but in a way that indicates they are not human. Though I do not know of any research that would confirm this, TMT studies demonstrate that the human-animal boundary is comforting when we are presented with reminders of our mortality. If animals do indeed make mortality salient to us, it would be helpful to have a quick way of reinforcing the human-animal boundary. The mental states we attribute animals—the ones resulting from a disgust-driven infrahumanization process—seem to play this role perfectly.

4.4 Ethical Implications of Disgust Toward Animals

It is clear from what I have said thus far that I think disgust research has normative implications for our treatment of animals. However, some ethicists are highly skeptical that disgust responses should have *any* bearing on normative questions, and would deny that the empirical research I have outlined is relevant to any substantive

problems in animal ethics. Dan Kelly (2011) refers to these ethicists as “disgust skeptics.” In short, disgust skeptics acknowledge that disgust responses place various limits on our moral judgments, but they do not think these limitations should have any role in normative ethics. Here I will attempt to refute “disgust skeptics.” First I will discuss arguments provided by Kelly and Martha Nussbaum, which state that disgust is not an epistemically reliable moral guide. Then I will discuss more generally the revisability of our disgust responses.

To be clear, I will not be arguing that moral disgust is a reliable guide to moral truth, nor will I argue that we should abandon the project of justifying our disgust responses. I am also not an advocate of the view, often criticized by disgust skeptics, that there is wisdom in repugnance (or that we inherit sound moral judgments through our disgust responses; Kass, 1997). Nonetheless, I think the disgust responses summarized thus far limit what ethicists can ask of people in their treatment of animals. To make my case, I will return to the discussion of psychological plausibility from section 2. Against Kelly and Nussbaum I argue that individuals’ disgust responses limit their moral obligation to animals. Here I draw specifically from the previous discussion of “ought implies can.” To supplement my argument against Kelly and Nussbaum, I also review research on the revisability of disgust, and suggest that the evidence further supports my claim that individual’s disgust responses limit their moral obligations.

4.4.1 Disgust Skepticism

Dan Kelly and Martha Nussbaum share the view that disgust is unreliable as a moral guide. They are opposed to those who think that deep moral wisdom is contained in our disgust responses (e.g., Kass, 1997). According to Kelly (2011), “widespread feelings of disgust are...simply irrelevant to the question of whether or not the norm itself is morally problematic or acceptable” (p. 149; also see Kelly and Morar, forthcoming). Nussbaum’s arguments against disgust are directed primarily at its role in law, but her conclusion is much the same as Kelly’s: “Disgust...offers limited guidance in a narrow set of laws concerned with physical distaste and danger. But when it becomes a constructive criterion of legally regulable conduct, and especially when it conduces to the political subordination and marginalization of vulnerable groups and people, disgust is a dangerous social sentiment” (Nussbaum, 2004, p. 171). Both are primarily concerned with human groups whose disgustingness has been cited as justification for differential treatment in law, politics, ethics, and society generally. The disgust some feel towards homosexuals, for instance, has been used as an explicit justification for differential treatment.

One could argue that disgust skepticism violates the minimum criteria for psychological plausibility. Disgust is pervasive in our moral lives, and any account of disgust must provide reasons to think we can manage our disgust in a way that avoids its negative consequences. However, Kelly and Nussbaum’s proposals on this topic are sufficiently insightful as to decrease the threat of total psychological implausibility. The bigger problem with disgust skepticism, which will be my focus, is that it is clearly

aimed at ethical prescriptions at the level of policy and law, and does not address what we can ask of individuals or groups of individuals. As discussed in section 2, ethical prescriptions at a societal level have the benefit of an abundance of resources, while the personal resources of individuals are much more limited.

To see the problem here, let's consider Nussbaum's view in more detail. Her main argument is that disgust has no role in a society that ascribes to the ideal that all persons are of equal worth. A strategy Nussbaum (2010) outlines in order to eradicate disgust and live up to this ideal is to adopt a "politics of humanity," where we all recognize the humanity we share with groups of people sometimes found disgusting (e.g. homosexuals). The most obvious problem with this is that it explicitly excludes nonhumans. However, we can imagine a "politics of animality," which focuses on similarities between humans and nonhumans. This might allow us to better control the disgust we feel towards animals. For instance, Nussbaum (2003, p. 423) urges societies to stop appealing to disgust in public policy and portraying entities as disgusting, even if that is in fact the primary emotion they evoke. These practices violate shared animality just as they violate shared humanity.

There are problems with this, however. First, and most obviously, Terror Management Theory predicts that highlighting shared animality will *increase* disgust felt towards animals. Though Nussbaum seems to recognize this problem, her solution is inadequate. For instance, drawing from Rozin's research, she suggests, "if the real issue underlying disgust is the fear and loathing people have for their animal bodies and their own mortality, then a society that wants to counteract its damages must go further,

addressing the body itself, and our anxieties about it” (2003, p. 424). Unfortunately, she offers no details on how societies are supposed to face this challenge. And even if she could do so, one wonders what this would entail for individuals. Mortality salience is thought to be deeply rooted in our psychologies. Without significant external support in overcoming the effects of mortality salience, individuals’ efforts would be totally insufficient. “Addressing the body itself,” as Nussbaum suggests, would only enhance the psychological threat.

Second, Nussbaum in fact rejects perfect equality between humans and nonhumans. Though she has argued at length that current inequalities between humans and nonhumans are unjustified, she nonetheless appears to think certain inequalities are justified, given fundamental metaphysical differences between humans and nonhumans (Nussbaum, 2004b, 2006). For example, in justifying certain types of animal experimentation, she states, “complex forms of life have more complex (good) capabilities to be blighted, so they can suffer more and different types of harms” (2006, p. 361). So a politics of animality would appear to be based on a false premise, even by Nussbaum’s own lights. One could perhaps argue that feeling disgust towards certain animals is justified, given these inequalities.

Assume for the sake of argument, however, that equality between humans and nonhumans becomes enshrined in law, and all disgust responses to animals have been eradicated from relevant laws and policies. This appears to be what disgust skeptics strive for. Even if this occurred, we cannot assume that disgust toward animals would be similarly eradicated from human psychologies. It is at least possible that human

psychology would remain unchanged, and thus individuals would face the same psychological constraints. On my account of psychological plausibility, as discussed in section 2, what would be asked of people legally, in such a case, would differ from what would be asked of them ethically. This follows from the previous discussion of ought implies can. If individuals cannot overcome their disgust responses, such that they cannot meet various moral obligations to animals, then they have no relevant moral obligations. An ethical theory suggesting otherwise would, on my account, be considered psychologically implausible.

In summary, my main objection to disgust skeptics is that their solutions are aimed at law and policy, and thus do not have implications for individual disgust responses. Of course, a disgust skeptic might reply with an account of how individual disgust responses are to be managed. In the next section, I will provide a sketch of such an account on behalf of disgust skeptics, aimed at the revisability of disgust.

4.4.2 Revisability of Disgust

Perhaps the key feature of disgust is that it is highly unmodifiable and automatic. Numerous experiments illustrate that disgust is extraordinarily resistant to change through reason (Russell & Giner-Sorolla, 2011; Shweder, Much, Mahapatra, & Park, 1997). Consider Jonathan Haidt's (2001) "moral dumbfounding" experiments. In these experiments, people were given vignettes describing acts that tended to elicit disgust and subsequent moral disapprobation but did not involve any harm. For example, one vignette described a case where a family eats its dead dog; another described eating a

dead chicken one has just had sex with. Participants in these experiments tended to say that these actions were morally wrong, but, when pressed by the researchers, they were incapable of cogently explaining *why* they were wrong—they were dumbfounded. The explanation Haidt offers is that moral judgments are driven by emotionally based intuitions, while other more reflective processes factor in later and at the margins of our moral judgments.

These results are relevant to the revisability of moral emotions, which many think is fundamental to moral judgment (e.g., neosentimentalists like D’Arms and Jacobson 2000, 2003, 2007 and McShane, 2007, 2011). Many ethicists do not deny the intransigence of emotions. For normative ethics, however, ethicists tend to emphasize the importance of the *attitude* we take towards our emotional responses. As D’Arms and Jacobson (2000) say, “to make an evaluative judgment is not to *have* but to *endorse* a sentiment” (their emphasis; p. 729). For instance, an observation many ethicists rely on is that we all experience moments where we disagree with our emotional responses. And, more importantly, this disagreement at least occasionally results in a change in our emotional responses, and with them our moral judgments. In this way moral inquiry seems to be characterized by its open-endedness, where whatever I currently am disposed to think is right or wrong is, in principle, revisable. Even if we know emotions are intransigent, the ethically right thing to do is to evaluate and attempt to change them if necessary.

But Haidt’s research indicates that the revisability of moral judgments does not entail *actual* revision; reflection often leaves the content unchanged (Haidt, 2010). That

is, even if it is possible for us to appraise our emotional and moral responses, this does not mean we actually do so with any regularity, or that, when we do, any revision actually ensues. Moreover, Haidt also contends that reflecting on one's emotions sometimes only reinforces their power and influence. Reflection commonly is used to justify our initial reactions, rather than work to revise them. As he explains the results of his dumbfounding experiments, "The refutation of...arguments does not cause people to change their minds; it only forces them to work harder to find replacement arguments" (Haidt, 2006, p. 218). So even at the margins of our moral lives, reason often fails to actually lead to revision in moral beliefs and emotions.

There have been various criticisms of Haidt's dumbfounding experiments, as well as of the more general claim that reason is ineffective in revising our emotions. A number of critics have argued that Haidt has overstated his claims about the role of automatic attitudes, and that in fact reason has a larger influence than he recognizes. For instance, in reviewing the empirical evidence for the revisability of emotions, Hanno Sauer claims:

Although emotions are often beyond our rational control, they can still be rationally amenable. This feature might also be called the reason-responsiveness of emotions. They are the proper object of evaluation and critical reflection, and we can—and often do—ask whether having a particular emotional response to a particular emotionally significant event makes sense. (2012b, p. 103)⁴⁴

⁴⁴ These sorts of claims are also frequently made by neosentimentalists. For instance, D'Arms (2005) says our emotions are "responsive to reason" (p. 10). D'Arms (2013, p. 6) claims that it is "not implausible" that we can develop control over our emotions. Other times (e.g., D'Arms and Jacobson, 2003, p. 144) neosentimentalists assert that we are capable of controlling our emotions to a sufficient degree, even while acknowledging that emotions are elicited automatically and frequently outside of conscious control.

I think most of the criticisms of Haidt's research are largely incapable of affecting the arguments I have presented here. For my purpose—responding to disgust skeptics—all that needs to be shown is that it is extremely difficult to control automatic attitudes, not that they are entirely unrevisable. Disgust, for instance, is “rationally amenable” in a way that matters only if individuals can have a reasonable chance to revise their disgust responses with the resources available to them. The research on disgust, however, especially that cited by Sauer and others (e.g., Kennett and Fine, 2009), fails to demonstrate this.

For instance, consider a study by Gallo, Keil, McCulloch, Rockstroh, and Gollwitzer (2009), which Sauer (2011) cites in support of the revisability of disgust. What the results of this experiment showed was that telling oneself “If I see blood, then I will stay calm and relaxed!” decreased disgust responses to pictures of blood, but “I will not get disgusted!” had no effect. An identical phenomenon was observed with fear of spiders. This is evidence that disgust is revisable, but only with significant support. Successfully modifying one's disgust response requires “implementation intentions,” or basic if-then commands that, at least in this experiment, *prepare* the participant for the oncoming stimulus and dictate a specific response. Desiring to change one's emotional response and even making a concerted effort to do so are insufficient. Without the right “coaching” or some form of properly informed external support, it is unlikely that individuals will be able to modify their disgust response. This experiment only proves Haidt's (and my) point that automatic attitudes are enormously difficult to revise.

I conclude that Haidt's research, and research from psychology more generally, indicates that moral emotions generally, and disgust specifically, are incredibly difficult to revise. Ethicists might be right that it is important to revise our emotional responses when morality demands it, but they need to provide some explanation for how ordinary human agents could eventually come to see this as well. The research I have summarized suggests that readily revising our moral emotions is not a prominent feature of our moral psychologies.

This provides a further response to disgust skeptics. Nussbaum's position in particular, as described above, largely avoids modifying individual disgust responses. Instead, she seeks to change the role of disgust in law and policy. However, I have suggested that individuals cannot revise their disgust responses by their own powers of reflection. If this is true, then individuals' moral obligations to animals would seem to be unaffected by Nussbaum's argument (and the argument of disgust skeptics more generally). Various moral duties to animals will require people to overcome their disgust responses, but without sufficient external support in doing this, people cannot be expected to meet those moral demands.

4.4.3 Concluding Thoughts

Ethicists should strive to provide normative prescriptions that are ethically acceptable while still acknowledging psychological limitations. The best evidence we have suggests that modifying our disgust responses to animals is incredibly difficult. Putting forth a theory that requires or presupposes that animals do not elicit disgust is

thus psychologically implausible. Urging universal compassion for animals or unqualified benevolence are the sorts of views that would seem to be excluded by my account. A general class of views that seem problematic, but would require more analysis than I can provide here, are those that adhere to what Garner (2013) calls the “equal consideration of interests principle” (as originally developed by Peter Singer). In short, these views require that the comparable interests of sentient beings deserve equal consideration, regardless of whether the beings in question are human or nonhuman. These views seem problematic because the disgust response, specifically when driven by TMT, is specifically attuned to whether something is human or nonhuman, and will influence one’s treatment of a being’s interests. In practice, therefore, people will be unlikely to abide by the equal consideration of interests principle. This is particularly problematic for views like that of Donaldson and Kymlicka (2011), where people are asked to accept animals as “co-citizens,” and treat animal interests roughly on par with the interests of other marginalized human citizens (like those with disabilities). My prediction would be that such an explicit comparison of humans to nonhumans would elicit disgust-based avoidance, leading people to reject this sort of proposal.

My main recommendation is for ethicists to focus on ethically acceptable options that exist within the constraints provided by disgust. Typically there will be many such options, though I leave it to ethicists to determine what these options are within their theories. By outlining the role of disgust in evaluating animals, I hope to have provided a way for ethicists to hone in on their psychologically plausible normative options.

5. REJECTING EMPATHY FOR ANIMAL ETHICS

A source of frustration for animal ethicists (as well as many animal advocates) is persisting apathy in the face of obvious moral harms done to animals. The remedy often recommended for this problem is to emphasize and promote *empathy* in human moral psychology.⁴⁵ For instance, we might think apathy results from a failure to understand what it feels like to be an animal under blighted conditions. Maybe dogs kept in kennels all day, pigs whipped and prodded on their way to slaughter, immobilized chickens in battery cages, monkeys forced to perform tricks for movie producers, and predatory animals confined to small glass-encased rooms for zoo audiences could all be avoided if people just understood what it felt like to be in such circumstances. Perhaps empathy can provide this sort of insight.⁴⁶

My argument here will be that empathy should not have a privileged role in our theorizing about animals. My argument is based on the rejection of the idea that empathy is *central* to moral concern. The centrality thesis, as I will call it, is a psychological claim: empathy is psychologically central to showing moral concern for animals. Moral concern, as I understand it, consists of being attentive to another's positive well-being.

⁴⁵ The role of empathy in moral judgment and action has seen something of a revival in recent years, among both philosophers and psychologists (e.g., Coplan & Goldie, 2011; Held, 2006; Howe, 2013; Oxley, 2011; Slote, 2007, 2010; Stueber, 2006, 2012). This has been the case in animal ethics as well. Notably, the increased emphasis on empathy in animal ethics has been quite rigorously tied to new empirical studies in cognitive science. Lori Gruen (2009, 2012) and Elisa Aaltola (2012), for instance, both draw heavily from the latest research in cognitive science to advance their own theories of how empathy improves moral treatment of animals. This stands in contrast to earlier writings on empathy and animal ethics that failed to draw on relevant empirical research (e.g., Donovan, 1996; Luke, 1995; Shapiro 1994).

⁴⁶ As Martha Nussbaum says, in commenting on a passage on empathy from J. M. Coetzee's novel *The Lives of Animals*, "[I]magining the lives of animals makes them real to us in a primary way, as potential subjects of justice" (Nussbaum, 2006a, p. 355).

This can be expressed in action or attitude. For instance, worrying that feral cats in my neighborhood are not having their basic needs met is an expression of moral concern, as is actively providing these cats with food and water. The basic claim of the centrality thesis is that empathy produces moral concern for animals better than any other emotion or psychological process. Ultimately I deny that that this is true; I reject the centrality thesis. I argue that other moral emotions, particularly moral anger, are more central to producing moral concern for animals, and thus more suitable than empathy for various normative aims in animal ethics.⁴⁷

To make my case against the centrality of empathy, I investigate six different empirical claims commonly made about empathy towards animals: (1) empathy predicts moral concern for animals (Aaltola, 2012; Donovan, 1996; Gruen, 2009; Luke, 1995; Sevillano, Aragonés, & Schultz, 2007; Shapiro, 1994), (2) high levels of moral concern for animals are caused by high levels of empathy (Gruen, 2004; Rachels, 2011; Signal & Taylor, 2007), (3) empathy towards animals causes empathy towards human beings (Ascione, 1992; Ascione, 2008; Munro, 2005), (4) cruelty towards animals is caused by a lack of empathy (Aaltola, 2012; Ascione, 1999; Fox & McLean, 2008), (5) empathy expands our moral concern to outgroups (Gruen, 2009; Gruen 2012; Solomon, 1999),

⁴⁷ My analysis is similar in nature to Jesse Prinz's (2011a, 2011b) arguments against the *necessity* of empathy for moral concern (particularly with respect to human beings). For instance, he asks, without empathy, will we will judge animals negatively and deny their needs? Without empathy, can we meet our moral duties towards animals? These questions provide the background to my investigation. I more directly address the question of centrality, which is slightly different. Even if empathy is not necessary for morality, we still might think that it is much more important than any other moral emotion. For instance, perhaps it is empathy that primarily determines positive evaluations of animals, even if other routes to positive evaluations exist as well.

and (6) empathy motivates us to act on our moral duties to animals (Aaltola, 2012; Berenguer, 2007). If the empirical evidence fails to support these claims, then it is not clear how empathy could be central to showing moral concern for animals. I find all six claims to be problematic, though some are more plausible than others.

First, however, I will attempt to get clear on what empathy means and which features of empathy are most relevant to showing moral concern for animals. The first section will take on this task. After laying out what I take empathy to be and how it functions with respect to animals, I will address each of the six empirical claims in turn. The final section will explore the implications of my arguments for animal ethics more broadly.

5.1 What is Empathy?

In this section, I provide a basic explanation of what empathy is and how it works, particularly with respect to nonhumans. First, I will outline emotion-sharing, which is the process I take to be essential to empathy, and the observations of emotional expressions and external conditions (which includes one's immediate surroundings and life circumstances) that can create shared emotions. Second, I will argue against the relative importance of cognitive empathy. Third, I will discuss the emotions that often result from empathy, and explain why they seem to pose significant challenges to the hypothesis that empathy is central to producing moral concern.

5.1.1 Emotion Sharing

The most crucial element of empathy is emotion-sharing. The functional aim of empathizing is to share in the emotions of another, or what is sometimes called vicarious emotion, emotional transfer, or affective resonance.⁴⁸ Though high-level cognition can influence this process in various ways, most psychologists and neuroscientists agree that empathy is largely automatic and unconscious. For instance, empathy is often described in terms of emotional contagion, or the rapid, unconscious sharing of emotions that results from observing another's emotional expressions (Hatfield, Rapson, & Li, 2009; Klimecki & Singer, 2013; Prinz, 2011b). Typical examples of emotional contagion are instances where we see someone smile or frown and cannot help but feel happy or sad ourselves. To be clear, empathy is not *reducible* to emotional contagion but is best characterized by the automatic emotion-sharing that results from the contagion process.

There are two main causes of the emotion-sharing provided by empathy. One, as just described, is observing another's emotional expressions. Emotions can be expressed through a variety of behaviors, but characteristically empathy results from observations

⁴⁸ The features of empathy I identify here do not pretend to fully capture any one person's views but rather draw from many prominent accounts of the basic processes involved in empathy. More comprehensive summaries of empathy can be found in Preston and deWaal (2002) and Preston and Hofelich (2012). A detailed definition of empathy from one of the field's leaders can be found in Batson (2009), in which he identifies eight prominent, but different, uses of the term empathy in the literature: 1. Knowing another person's internal states, including his or her thoughts and feelings. 2. Adopting the posture or matching the neural responses of an observed other. 3. Coming to feel as another person feels. 4. Intuiting or projecting oneself into another's situation. 5. Imagining how another is thinking and feeling. 6. Imagining how one would think and feel in the other's place. 7. Feeling distress at witnessing another person's suffering. 8. Feeling for another person who is suffering.

of the face. The second is observing another's *external conditions*.⁴⁹ This includes another's immediate surroundings as well as their presumed life circumstances. For instance, I can empathize with a homeless man by observing where he sleeps or I can empathize by considering what his daily activities likely consist of (e.g. where he goes to get food).

Of course, observing another's emotional expressions or external conditions will not always result in emotion-sharing. To see the implications of this, we need to distinguish the *process* of empathy from the *results* of empathy. Emotion-sharing best characterizes the process of empathy. The intentional object of empathy is to share in the emotions of another, even if this fails to actually result in emotion-sharing. We can see this by looking at cases where people empathize with entities that do not plausibly even have emotions, like trees, ecosystems, and mountains (Schultz, 2000). In such cases, people show increased concern for the entities in question and claim to feel a change in emotion (Tam, 2013).⁵⁰ There are no success conditions based on the accuracy of the results here; it is nonsensical to ask whether people are actually sharing in the emotions of these entities. Yet clearly people are engaging in the process of empathy.⁵¹

⁴⁹ It is generally acknowledged by empathy theorists that understanding the particulars of another's situation is crucial to empathy (Maibom, 2012).

⁵⁰ Another interesting and illustrative case is the practice of using rubber rings to remove tails from cows. It turns out that cows do not feel any pain from this practice, according to behavioral and physiological measures (Tom, Duncan, Widowski, Bateman, & Leslie, 2002). However, if I am empathizing with a tail-docked cow, and have turned my attention to its circumstances, presumably I will feel pain, sadness, and other emotions appropriate to tail docking.

⁵¹ If brain images could be compared, for instance, of someone in the process of empathizing with a mountain and someone empathizing with another human being, and the images indicated identical emotion activation, I don't see any reason why we wouldn't call both processes true empathy.

Rather, the success condition of empathy is simply the elicitation of emotions in response to aiming to share the emotions of another. These are the results of empathy.⁵² Again, these need not result from conscious activities. Emotional contagion, for instance, sometimes elicits emotional activation that is incongruent with the target's emotions (e.g., if I confuse tears of joy with tears of sadness), thus failing to result in true emotion-sharing. Similarly, if we turn our attention to animals' living conditions (e.g., a pig's farrowing crate), and our empathy system is engaged, we should experience, emotionally, what it would be like to live in those circumstances.⁵³ The emotions we experience may bear no resemblance to what the pig is actually feeling, and thus do not constitute emotion-sharing, but we are nonetheless empathizing.

The account of empathy I favor, and have described thus far, is more capacious than that offered by many other empathy theorists in that I include a wider variety of psychological processes. I see this as a benefit of my account: the more capacious my conception of empathy the more interesting my rejection of the centrality thesis becomes. A traditional challenge in providing an account of empathy towards animals is determining whether we are truly sharing their emotions or instead projecting our emotions onto them. Most accounts of empathy in the case of humans address this

⁵² This still allows us to exclude explicit *projection* of one's emotions onto another being from the definition of empathy, which many empathy theorists agree is necessary in a proper definition of empathy (Coplan, 2011; Darwall, 1998; Goldman, 1993). Projection does not satisfy the requirements for the process of empathy.

⁵³ Jesse Prinz's (2011a) definition of empathy aptly captures the process of empathy, as I understand it: "Empathy is...a kind of associative inference from observed or imagined expressions of emotion or external conditions that are known from experience to bring emotions about" (p. 215).

problem by limiting legitimate empathy to *accurate* sharing of emotions, where this excludes self-projection. However, because it is difficult to provide success conditions for this process, empathizing with animals appears to be a deficient form of the real thing—certainly not something capable of providing moral insights.

My more capacious account, by contrast, takes into account a wide variety of emotional understandings that arise out of observing animal expressions and living conditions. I see empathy towards animals as legitimate empathy, worthy of the attention of ethicists, and want to include all evidence that might speak in favor of its centrality in producing moral concern. This makes my task more difficult, but if I am successful the implications are also thereby more significant.

5.1.2 Cognitive Empathy

Most studies of empathy do not test emotion sharing as such. Rather, they study what is called cognitive empathy, or empathy that comes about from actively imagining what another person is feeling. For example, the empathy prompt in much of Daniel Batson's classic work on empathy (which will be discussed in more detail below) is phrased, "try to imagine how [the target of empathy] feels about what has happened and how it has affected [the target's] life." This encompasses my definition of empathy but also requires a great deal of cognitive effort to successively achieve emotion sharing.

Some empathy theorists claim that cognitive empathy is the only true form of empathy. Many separate cognitive empathy from more basic emotion-sharing processes, like emotional contagion, because it is only with cognitive effort that we are able to

overcome various limitations in empathizing with others (e.g., in overcoming our self-focused bias; Coplan, 2011). What many philosophers have focused on is not the sharing of emotions, as I have emphasized, but rather the process of thinking about a person's situation, as well as what that person is like (including their psychological makeup), with the goal of recreating what that person likely is feeling.⁵⁴ Lori Gruen (2009), among animal ethicists, requires similar features for the process of empathy, "To accurately empathize, one must focus on and take account the specific context of the other, their idiosyncratic desires, and the processes that shaped those desires, their developmental and, in the case of non-sentient beings, their ecological and evolutionary histories, and their distinctive telos" (p. 33).⁵⁵

While the cognitive aspect of empathy is important, I think it has been overstated. Effortful empathizing enhances emotion-sharing, and can correct for certain biases, but it does not constitute the empathic process. It is not *required* that we engage in correction of what we think others are feeling in order to empathize. Rather, empathy only requires that we direct our attention to others' emotional expressions and living conditions, and feel new emotions as a result. I do not have to actively imagine the emotions of a pig in a farrowing crate. I just have to direct my attention to the pig being

⁵⁴ Stueber (2006) calls this process "reenactive empathy," Goldman (2011) calls it "reconstructive empathy," and Gruen (2009), discussing animals specifically, variously calls it "entangled" or "engaged" empathy. Ickes (1997) expresses this idea from a psychologist's point of view.

⁵⁵ Donovan (2006) makes a similar suggestion: "the originary emotional empathetic response must be supplemented with an ethical and political perspective (acquired through training and education) that enables the human to analyze the situation critically so as to determine who is responsible for the animal suffering and how that suffering may best be alleviated" (p. 323).

in a farrowing crate.⁵⁶ Empathy is capable of providing us with an understanding of another's emotions without further cognitive effort.

I see cognitive empathy as mostly a failsafe—a way of enhancing the salience of others' emotions when normal empathy processes have failed. With respect to animals, cognitive empathy is important if people fail to consider the plight of animals spontaneously.⁵⁷ For instance, if the automatic processes of emotional contagion fail to produce moral concern for animals in pain or in blighted conditions, people can actively imagine themselves into those conditions in order to elicit empathy.⁵⁸ At its foundations, empathy leads us to associate certain living conditions with the way it would feel to live under those circumstances, and to associate certain emotional expressions with the way it would feel to express those emotions. Perhaps we use cognitive effort if these activities fail to elicit emotion-sharing, but this is extraneous to the core processes of

⁵⁶ Though I won't address this in detail, my position on cognitive empathy helps avoid an objection raised by Holton and Langton (1999) against Peter Singer's use of empathy: "We suggest that if you want to discover the preferences of a bat or a platypus, the last thing you should do is to try to imagine yourself into their place. Is it cruel to screech loudly at a bat? Is it cruel to pass a 12 volt shock through the water near a platypus? We don't know, and imagining ourselves in their place won't help. Imagination 'yields a perfect blank' (p. 225). Empathy is not as problematic as they make it seem, because it is not the function of empathy to accurately discover the preferences of animals. It is merely to produce emotions in us by turning our attention to animals' behavioral expressions and living circumstances.

⁵⁷ As Hare (1963, p. 224) states in a provocative passage, "Those who indulged in bear-baiting did not reason: 'If we were bears we should suffer horribly if treated thus; therefore we cannot say that it is all right to treat bears thus.' The bear-baiter does not really imagine what it is like to be a bear. If he did, he would think and act differently."

⁵⁸ To see other popular definitions of empathy used by those writing about animals see Signal and Taylor (2007, p. 126) and Eisenberg (1988). I take these other definitions to be broadly consistent with my own definition.

empathy.⁵⁹ Importantly, there's good evidence that even high-level empathizing, as used in cognitive empathy, makes use of the same neural substrate as low-level empathizing, like that observed in emotional contagion (Iacoboni, 2011; Preston & Hofelich, 2012).

Another problem with emphasizing only the cognitive aspects of empathy is that doing so makes it difficult to separate empathy from sympathy. A common way to distinguish empathy and sympathy is to say that concern for another is constitutive of sympathy while it is not for empathy (de Vignement & Singer, 2006). The problem this raises is that while cognitive empathy does not require concern as such, it does require motivation and cognitive effort, which must come from somewhere. Unlike emotional contagion, cognitive empathy seems to require one to take enough interest in another's emotions to reconstruct what they are feeling. Following a point made by Prinz (2011a), if our inquiry is directed at whether empathy is necessary for expressing concern for animals, concern can't also be included in the definition of empathy, prior to the investigation. By making cognitive empathy the focus, we lose out on some of the benefits of emotional contagion while blurring the line between empathy and sympathy.

5.1.3 Resulting Emotions and Alternatives

The emotions produced by empathy are of particular importance for my investigation. Some of these are related to moral concern, and might seem to support the

⁵⁹ Elisa Aaltola (2012), among animal ethicists, similarly argues that emotion-sharing need not be accurate, but, unlike me, thinks empathy is primarily cognitive. As she says, "we can feel empathy towards non-human suffering if we use imagination to perceive what that suffering might be like, and if we feel care towards that suffering; yet we need not actually share the animals' experiences" (p. 166).

centrality thesis. Sympathy, for instance, can result from the emotion-sharing of empathy, which might help explain why empathy would be central to moral concern (if it is). Other emotions that result from empathy are problematic, however, and illustrate limitations for empathy's role in producing moral concern.

For instance, empathy often elicits emotions that decrease moral motivation. Take a simple example: you observe your neighbor beating his dog. If you empathize with the dog, the emotion-sharing component of empathy predicts that you will feel a twinge of pain. Likewise, if the dog is cowering in fear, it is predicted that you will share in the feeling of fear.

These responses are problematic because, as discussed in the previous section, these emotions would be categorized as *avoidance* emotions—emotions that generally motivate us to move away from a target or goal. If the goal of empathy is to produce helping behavior, then feeling emotions of pain or fear (or perhaps defenselessness) are not going to be very effective. Fear, for example, functions to remove ourselves from a threat; perceiving others in pain elicits an avoidance response, at least partly because pain signals that a threat exists in the immediate environment (Yamada & Decety, 2009); and sadness, another emotion we might feel in response to animal harm, similarly functions to reduce moral motivation (Carver, 2001). Feeling any of these emotions as a result of empathizing with animals in pain might produce morally commendable

behaviors in certain situations, but it is not likely to motivate us to intervene in cases of abuse.⁶⁰

There are other emotional responses that can be elicited simply by observing a situation. These emotions can include empathy, and sometimes result from empathy, but also arise directly from the observed circumstances. These raise problems for the centrality thesis because they are more effective than empathy, as such, in producing moral motivation.

Consider the case of moral anger. There is evidence that anger can arise out of empathy (Batson et al., 2007), but it also operates independently. When people perceive a threat of personal harm, they often respond with anger (Rozin, Lowery, Imada, & Haidt, 1999). For instance, a beating doesn't just produce pain and suffering, it also produces a response to stop the beating—namely, defensive anger. This is true even when observing as a third party. For example, observing one's friends receive threats of violence can elicit just as much, if not more, anger than if the threat is made to oneself. This might be because we empathize with the situation, but it is likely also due to the perceived *transgression*. The mere fact that someone wishes to inflict pain on another can be enough to elicit anger, especially if this is done unjustly. Anger is the converse of emotions like pain and fear in that it is categorized as an *approach* emotion (Carver &

⁶⁰ An argument someone might make against me here is as follows: In cases of emotion-sharing, our understanding of self and other allows the empathizer to realize that his or her emotions are different in important ways from the emotions of the target of empathy, and thus we are not susceptible to or can avoid the negative consequences I mention. However, many studies have shown that the emotion-sharing of empathy is quite literal—we feel others' emotions as if we were in their place (e.g., Singer, Seymour, O'Doherty, Stephan, Dolan, & Frith, 2004). Even when we acknowledge a strong self-other distinction, empathy often leads to congruence in emotion—including motivational direction—between the target and the empathizer.

Harmon-Jones, 2009; Harmon-Jones, 2003; Harmon-Jones, Gable, & Peterson, 2010).

Anger is motivating and helps us to achieve goals, especially in cases of personal harm.

This provides ethicists a reason to privilege moral anger, and not empathy.

Moral anger, and other emotions, will be returned to throughout this section.

These emotions will be used to make a firmer case against the centrality of empathy in producing moral concern. I turn now to the various claims made by the centrality thesis.

5.2 Empirical Claims of the Centrality Thesis

The six claims I will explore are meant to capture the most perspicuous empirical hypotheses about how empathy is related to expressing moral concern to animals. I will reject the idea that empathy is central to morality, but the argument will not come easily. There are indeed some convincing experimental demonstrations of the role of empathy in moral judgment. My main argument is that in many of these experiments we can identify other non-empathic processes that better explain the relevant phenomena. I will further draw from the point just made that other emotions appear to be more central to our moral response to animals (such as anger at transgressions against animals).

5.2.1 Empathy Predicts Moral Concern Toward Animals

If empathy is central to expressing moral concern towards animals, then at the very least we should find a *positive statistical correlation* between empathy and moral concern for animals. This should be the case either for dispositional empathy (as an enduring trait) or occurrent empathy (as elicited by experimental factors). Dispositional

empathy, for instance, does indeed predict helping behaviors towards other human beings (Davis et al., 1999), so perhaps a similar relationship exists with animals. I think that such a relationship *does* exist, but it is not strong enough to support the centrality thesis.

Consider first an experiment by Sevillano, Aragonés, and Schultz (2007), whose findings are characteristic of other research on empathy and attitudes to animals (which will be discussed throughout this section). They measured both dispositional and occurrent empathy. Like much of the classic research on empathy (e.g., that of Daniel Batson), participants received one of two prompts to complete their task. They were asked either to remain neutral and objective or to take the target's perspective and to think about how the target feels. The latter prompt is the most widely used method of testing the effects of empathy, and will come up again.

Participants were shown ten photos of animals, five of which depicted animals being harmed (e.g., a seal caught in a fishing net) and five just of animals in nature (e.g., a rhinoceros on a savannah). Participants' level of concern for the environment was also measured.⁶¹ Results showed that neither dispositional empathy nor the empathy prompt, taken on their own, correlated with concern for the environment. However, when those who received the empathy prompt *also* saw the photos of animals being harmed, they *did*

⁶¹ Concern for the environment was measured using Schultz's (2000) Environmental Motives Scale. The prompt for this scale states, "People around the world are generally concerned about environmental problems because of the consequences that result from harming nature. However, people differ in the consequences that concern them the most. Please rate the following items from 1 (not important) to 7 (supreme importance) in response to the question: I am concerned about environmental problems because of the consequences for [blank]." Participants then choose among a range of entities that might be affected by environmental problems.

express increased concern for the environment (which is also consistent with previous studies using similar stimuli and only induced empathy; Schultz, 2000). There were no effects observed for dispositional empathy. This suggests there is a relationship between perceiving animal harm, being induced to feel empathy, and expressing concern for the environment.

There is one caveat to this experiment, however, one which pops up repeatedly in empathy experiments. Those who did not receive the empathy inducing instructions, but who were instead told to remain objective when viewing the pictures, also scored relatively high (to statistical significance) on concern for the environment after viewing animals being harmed.⁶² These participants still scored lower than those who received the empathy prompt, but this indicates that empathy is not *necessary* for moral concern, just that it is slightly more effective than being told to remain objective.

Of course, Sevillano et al.'s (2007) study measured environmental concern, not concern for animals as such. But other research more directly on empathy and animals has provided a similarly complicated and ambiguous picture. For instance, consider one of the most widely cited pieces of evidence for the importance of empathy, from an experiment conducted by Shelton and Rogers (1981). They showed people films of whaling activities, some of which were higher in intensity (e.g., more gore) than others. Participants in one condition were given no special instructions while watching the films (low empathy), while those in another condition (high empathy) were told to try to feel

⁶² For those interested in the details, the empathy condition was only slightly higher ($b = .541$, $t(181) = 2.496$, $p = .02$) than the objective condition ($b = .631$, $t(181) = 2.006$, $p = .05$).

what the whales were feeling, and to sympathize and empathize with the whales (that sympathy was included in the prompt is problematic methodologically, but that will not matter for the rest of the discussion). After the films had been played, participants filled out an emotion scale, designed to measure the effect of the film on their emotional state, and were also asked if they would be willing to provide assistance to whales in the future.

The results showed that levels of empathy increased after viewing the high-intensity films, regardless of whether people received the empathy prompt (that is, for participants in both the low and high empathy conditions). But only those who received the empathy prompt showed increased empathy after either high- *or* low-intensity scenes. So intense harm elicited empathic responses not elicited by low-intensity harm, but both types of harm elicited additional empathic responses when combined with the empathy prompt. Those in the high-empathy condition were also more likely to declare an intention to help whales. Overall, this evidence suggests that something about the empathy induction led to greater moral concern for whales.

This study is important because of the fact that low-intensity harm produced a significantly lower empathic response as well as a decreased intention to help in people who had not received the empathy prompt. This appears to illustrate that empathy plays an important role in responding to harm. The whales in the low-intensity film were clearly being harmed, but because the harm was not made salient, people were not induced to provide assistance. Besides indicating the importance of empathy generally, the results might also be taken to indicate that cognitive empathy and perspective taking

are important safeguards in our moral thinking, capable of correcting us when our natural sensitivity to animal pain fails to elicit helping behavior.

However, it's not clear from the experiment whether those who did not receive the empathy prompt were truly *without* moral concern. Those in the low empathy condition—where, remember, they received no instructions at all—still scored relatively high on their willingness to help. The overall means, on a ten-point scale for willingness to help, was 7.8 for those who received the prompt and 6.8 for those who received no instructions. When viewing the low intensity films specifically, those who received no instructions had a mean of 5.5. Remember that levels of reported empathy before and after the films did not change for participants in this condition. While 5.5 is a fair distance from the average score of 8.3 for those in the high intensity, high empathy condition, it's certainly not zero. So it's not the case that those in the low empathy condition were blind to the harm caused by whaling, just that they would be much more willing to help if they had received the empathy prompt.

I take a measured conclusion from the two experiments just described. Empathy amplifies our moral concern for nonhumans, but it is not central to this moral concern, and certainly is not necessary. Participants in both experiments expressed relatively high levels of concern for nonhumans prior to any empathy induction. In Sevillano et al., this preexisting concern had no relationship to dispositional empathy, and in Shelton and Rogers, one group of participants expressed a significant degree of concern without any measurable change in empathy.

Shelton and Rogers' experiment also highlights the role of other, non-empathic, processes, in producing moral concern. Consider how moral anger likely plays a role in people's responses to the whale videos. An alternative explanation for their results is that asking participants to empathize with the whales caused anger, which then motivated people's willingness to help whales. This is consistent with the fact that anger is an approach-related emotion, and is also consistent with other studies on the approach-motivation of anger. Harmon-Jones et al. (2003), for example, found that college students who were angry about increased tuition costs showed increased activation in approach-motivational areas of the brain when told they would have an opportunity to sign a petition to prevent future increases in tuition. That is, having the opportunity to perform some action directly related to one's anger actually increased approach-motivation. Relatedly, Tagar, Federico, & Halperin (2011) found that people who were more angry about a political conflict were more likely to propose solutions when asked how they would ameliorate the conflict (also see Batson et al., 2007).

These studies illustrate the motivational strength of anger. Harmon-Jones et al.'s experiment in particular illustrates that approach-motivation produces carryover effects—the sort that could explain the willingness to help in Shelton and Rogers' experiment. As explained above, the intensity of the whale films actually had a greater effect on the declared intention to help whales than did the empathy induction. This is unsurprising when we consider that harm, especially unjustified harm as it was depicted in the film, normally elicits outrage among Westerners. Though empathy might play a role in this, the more central moral emotion would be anger.

The study showing the strongest positive correlation between empathy and concern for animals is Henry (2006). Using the Attitudes Toward the Treatment of Animals Scale and the Interpersonal Reactivity Index (common assessments of attitudes towards animals and dispositional empathy), Henry found a 0.43 correlation between dispositional empathy and caregiving attitudes toward animals. These results certainly are promising and can be taken as evidence that empathy is positively related to moral concern for animals.

However, again there are caveats. Numerous other studies have found much lower correlations. For instance, Taylor and Signal (2005) found a 0.33 correlation between empathic concern and positive attitudes toward animals among Australians, and Erlanger and Tsytarev (2012) found a 0.32 relationship between dispositional empathy and discomfort toward animal cruelty (for similar data on Norwegians see Ellingsen, Zanella, Bjerkås, and Indrebø, 2010). These lower correlations are significant for a couple of reasons. One is that a common heuristic used by personality psychologists is to treat as unreliable any correlations lower than 0.30 between traits and behaviors. Such a low correlation cannot be relied upon for accurate prediction. At this level, any correlation observed between a trait and a behavior in one instance does not reliably predict that correlation in any other instance. All of these correlations are above the 0.30 threshold, but just barely. This is particularly important in light of a second consideration, which is that these studies rarely test multiple variables, and when they do, they tend to find multiple significant positive relationships. To take one example, Erlanger and Tsytarev also found that perspective taking and personal distress,

considered separately from empathy, were significantly correlated with discomfort toward animal cruelty. The correlation with these factors was relatively small (0.15 and 0.16 respectively), but the weak correlation with empathy makes it difficult to say it was primarily empathy, and not a collection of related variables, that was responsible for the observed discomfort.

The weak relationship found in the majority of these studies should make one skeptical of the centrality of empathy. The only strong conclusion that can be taken from these studies is that there is a slight positive correlation between empathy and concern for animals. There is also good reason to suspect that other emotions, like anger, are more central.

5.2.2 Those Who Express the Most Concern for Animals Do So Out of Empathy

The experiments just described could be interpreted as suggesting that, at the very least, possessing a *great* capacity for empathy is a good thing—animals will be treated better by those who are more empathic. This can be phrased as an explanation for the phenomenon of animal loving and animal protectionism: high levels of empathy cause increased concern for the welfare of animals. If this is true, then ethicists have a good reason to promote empathy.

Researchers have indeed found some support for a positive relationship between empathy and abnormally high concern for animals. Signal and Taylor (2007), for instance, found evidence for this relationship among self-identified animal protectionists (e.g., people who work at animal shelters). One interesting feature of Signal and Taylor's

results is that empathy, but not perspective taking, was correlated with positive attitudes toward animals. As explained above, it is important to distinguish between empathy as automatically sharing emotions with another (like emotional contagion) and empathy as perspective taking. Animal protectionists generally scored high on both empathy and positive attitudes toward animals, but this was apparently not something that was related to actively taking the perspective of animals. This is interesting because it suggests that empathizing with animals is an automatic attitude among this group of people, and not something that has been taught or is under their direct control.

There are other complications with interpreting this data, however. First, while there is good evidence that being for or against animal experimentation correlates with differences in empathy (Broida, Tingley, Kimball, & Miele, 1993; Furnham, McManus, & Scott, 2003), it's not clear what causal role empathy plays. For instance, Furnham et al. found that opposition to animal experimentation also correlated with the personality traits of openness, agreeableness, and introversion, which may serve as more fundamental explanations of pro-animal attitudes.

Other evidence on this relationship is ambiguous and somewhat inconsistent with Signal and Taylor's results. Plous (1993) found that when viewing pictures of animals in pain, self-described animal activists ascribed more pain to the animals than did non-activists. However, in a separate measure of skin conductance, which measures physiological response, there were no significant differences between activists and non-activists. If differences in pain ascription led to greater discomfort among activists, or were caused by shared emotions (i.e., through empathy), there should have been

corresponding differences in the skin conductance measure. The automaticity of empathy, as suggested by Signal and Taylor's (2007) evidence, would predict that skin conductance would bear this out.

Some have suggested that vegetarianism is due to enhanced empathy (Gruen, 2004; Rachels, 2011), which could be construed as an argument that increased empathy causes vegetarians to care more about animals. While some studies have found no differences in empathy between vegetarians and non-vegetarians (Preylo & Arikawa, 2008), others have found evidence to suggest the contrary. I will briefly describe an experiment by Filippi et al. (2010), which some have claimed shows a connection between empathy and vegetarianism.

Filippi et al.'s experiment showed scenes depicting torture of humans and animals to self-described vegetarians, vegans, and omnivores. Results showed more activation in areas of the brain that process emotion in vegetarians and vegans than in omnivores when torture scenes were viewed. This seems to suggest that eating practices reflect a difference in emotional processing, which one could argue is due to differences in empathy.

I again think there are other more plausible, non-empathic, explanations for these results. One is that these differences in emotion processing might be a result of differential responses to moral transgressions. Vegetarians and vegans are likely to have strong moral convictions. These would likely be activated by torture (particularly of animals), thus producing greater emotional responses. Thinking that torture is wrong is sufficient to produce a strong emotional response when presented with graphic evidence

of torture. Another possible explanation is that vegetarians and omnivores responded differently to specific features of the torture scenes. For instance, some of the pictures showed mutilated bodies, which tend to elicit disgust and fear, with no further evidence to suggest torture specifically. Some studies have found that disgust sensitivity in particular is known to predict vegetarianism (Rozin, Markwith, & Stoess, 1997). So perhaps the role of disgusting stimuli explains the difference in emotion processing. Omnivores still expressed moral activation in response to scenes of torture, indicating that they responded emotionally, just not in the same way as vegetarians and vegans.

From these studies, I conclude that the evidence is, at best, ambiguous as to whether high rates of empathy produce greater moral concern for animals. One final complication I'll mention is that greater empathy towards animals appears to come at a cost. Knight, Bard, Vrij, and Brandon (2010) found that animal welfarists showed more empathy towards animals and less empathy towards human beings. Animal welfarists' enhanced empathy towards animals, that is, was associated with a corresponding decrease in empathy towards humans. Among other groups, like scientists, empathy toward animals was positively correlated with empathy toward human beings, but the overall level of empathy to either group was much lower than that expressed toward animals by animal welfarists. This research did not establish a causal relationship, but it does suggest a troubling possibility: our capacity for empathy is limited, and high degrees of empathy only result from narrow focus on a single moral object. Empathy might assist and support moral concern for animals, but only in instances of single-minded devotedness to their causes, to the exclusion of others (a finding that would be

consistent with some research on human empathy as well; Batson, Klein, Highberger, & Shaw, 1995).

5.2.3 Empathy Towards Animals Increases Empathy Towards Humans

The “sentinel hypothesis” (Patterson-Kane & Piper, 2009), or what is often referred to as “The Link” (Linzey, 2009), claims that cruelty to animals indicates a corresponding propensity for cruelty towards humans. Kant and Aquinas both famously advocated this position. An apparent corollary of this is that acting kindly towards animals indicates a kindness towards humans. We can hypothesize that the relationship runs in the opposite direction as well: cruelty and kindness toward humans indicates a similar propensity toward animals.⁶³

This family of relationships is important for empathy because it indicates that the capacity for empathizing, or sharing emotions with entities other than oneself, is generalizable. We can call this the *generalization hypothesis*: kindness (or cruelty) to one category of beings (humans or nonhumans) will increase kindness (or cruelty) to the other. This section will investigate this hypothesis by specifically examining the role of empathy.⁶⁴

⁶³ Some of the first research on how attitudes towards one group cross over to the other can be found in Levinson (1978).

⁶⁴ One example of people advocating something like the generalization hypothesis comes from Munro (2005, p. 66). She reports that self-described animal protectionists often ask people to apply the empathy they show towards their pets and other human beings to a wider variety of animals.

The Knight et al. (2010) study just discussed identifies two potential implications for the generalization hypothesis. On the one hand, self-described animal welfarists seem to refute the generalization hypothesis: their empathy is attuned more to animals than to human beings. But scientists, on the other hand, showed the exact relationship predicted by the generalization hypothesis (as did other groups). One main difference between these groups of people, we might think, is their experiences with animals. Usually The Link is discussed in terms of how early childhood experiences with animal abuse teaches children about abuse more generally. The generalization hypothesis can be taken to make a similar claim, such that empathizing with animals teaches people about empathy more generally.

In interventionist situations, where animals are introduced explicitly in order to change attitudes towards human beings, it has been found that children's empathy towards humans does indeed increase as a result of interacting with and learning more about animals (Ascione, 1992; Thompson & Gullone, 2003). However, in these studies children learn about much more besides empathy. In Ascione (1992), for instance, children were also taught about justice and pain and how these concepts applied to animals. This makes it difficult to conclude that empathy is responsible for the observed increase in empathy towards humans.

This raises a question for the generalization hypothesis: what type of animal experience is necessary for the cultivation of empathy? Mere exposure is insufficient, but more extensive training will cultivate much else besides empathy. One plausible proposal is that what is needed is providing care for animals—ensuring that their basic

needs are met. This would plausibly also teach people how to provide care for other human beings, and the mechanism responsible would be empathy.

One piece of evidence against this care feature is that self-described animal protectionists generally score high on empathy while farmers generally score low (Hills, 1993). Farmers, arguably, spend a lot of time providing care for animals, so they should possess more empathy. Of course, someone could argue that this is unsurprising because farmers exploit animals and use them for human ends. Perhaps, someone might say, cultivating empathy requires the *right kind* of care, the kind farmers generally do not engage in. But it seems unlikely that farmers do not possess the right experiences. It is often pointed out that farmers are quite sensitive to the welfare of their animals, as, at the very least, they must be in order to gain a profit and meet contemporary agricultural standards. Thus, farmers should show great empathy if caring for animals possesses a strong relationship to empathy.

Other empirical research can help sort out these issues. Evidence that runs contrary to the predictions of the generalization hypothesis can be gathered from Paul (2000). Paul found that, among Scottish participants, the correlation between empathy towards humans and empathy towards animals was significant but still below the .30 cutoff discussed above (.26). More problematically, the results showed that owning a pet correlated with animal empathy but *not* human empathy. So the connection observed between empathy towards animals and empathy towards humans, slight as it was, did not come about because of interactions with animals (or at least not with pets).

Paul's (2000) results point to a common phenomenon observed with empathy: we tend to be partial, and empathize with those who have special relationships with us. Pet owners, for instance, show more empathy towards their pets than towards human beings. This will be discussed in more detail in later sections, but its role here is important because it functions as a counter to the generalization hypothesis. Rather than generalize our empathy from animals to humans, it is instead the case that we empathize with those we are close to, and this often involves both animals and humans.

To further illustrate this, consider a set of experiments conducted by Angantyr, Eklund, and Hansen (2011). They gave people in Sweden a story describing a situation in which a man, woman, cat, or dog had been found lying on a street with broken ribs and a punctured lung. Participants were then asked to rate 16 emotions they were feeling towards the target, including empathy. In other conditions, the same situation was described but for a human child, human infant, and a puppy. Interestingly, they found that parents showed more empathy for infants and pet-owners showed more empathy for puppies. This difference did not show up for the adult pets or for the human child. The researchers did find that empathy to animals (either dog or cat) correlated positively with empathy to humans (either male or female), but experience with puppies and infants created additional, localized empathy, particular just to those groups. The generalization hypothesis would predict the opposite, that any localized empathy would generalize to others.⁶⁵

⁶⁵ A study from Batson, Lishner, Cook, and Sawyer (2005) might be cited in favor of the generalization hypothesis. They attempted to elicit biased (or partial) empathy by asking college

I interpret these results as speaking against the generalization thesis insofar as there is no discernable causal connection between empathy towards animals and empathy towards human beings. Rather, we have enough positive relationships with both animals and humans that we naturally respond to either group with empathy. In the studies just discussed, neither empathy towards puppies nor empathy towards infants showed any generalized effects. Pet owning also failed to transfer to empathy towards humans. More direct tests are needed, however. For instance, if someone low in dispositional empathy were to be introduced to animals, and subsequently increased their empathy towards both humans and animals, that would provide stronger evidence for the causal role of empathy.

5.2.4 Cruelty to Animals is Caused by a Lack of Empathy

The idea that empathy is necessary for moral concern entails the further claim that without empathy we cannot and will not express moral concern. Some take this further and say that a lack of empathy causes cruelty (Ascione, 1999; Baron-Cohen, 2011). We should of course promote empathy if doing so will avoid the creation of cruelty and evil.⁶⁶

students to read about another college student who had broken her leg and was struggling through rehabilitation. Instead, they showed equal empathy to the student, a 5-year-old pet dog, a 3-year-old human, and a 4-month-old puppy. However, I find it unlikely that college students would be partial to other college students, and thus this study does not provide results relevant to the generalization hypothesis.

⁶⁶ Here I understand “cruelty” to encompass both active (e.g. intentional abuse) and passive (e.g. neglect) forms of mistreatment. This is generally how the term is used in clinical (Ascione, 2008) and legal (Favre, 2011) contexts.

One way of being cruel to animals is by abusing them. Some have claimed that only by urging people to empathize can animal abuse be avoided (e.g., Aaltola, 2012, p. 160; Fox & McLean, 2008). However, evidence linking abuse of animals to lack of empathy is mixed. Henry (2006), for instance, surveyed people on their past experiences with animals, including past caregiving as well as past abuse, in addition to measuring participants' empathy. Results showed that empathy was a strong predictor of caregiving attitudes toward animals regardless of past experiences with animals. Importantly, those who reported participating in animal abuse were no different than others on the empathy measure. Whatever caused people to participate in animal abuse, it was not a failure of dispositional empathy. In fact, the only predictor of animal abuse was having been sexually abused as a child.⁶⁷

A frequently cited counterexample to the cruelty-empathy link is the widespread love of animals in the Third Reich. The Nazis notoriously proposed and passed ambitious animal protectionist laws, advocating for anti-vivisection, humane hunting, pain reduction in livestock animals, and species preservation, as well as promoting vegetarianism as a moral ideal (Arluke & Sanders, 1996; Sax, 2000). Moreover, they often justified these ambitious laws by appealing to the non-instrumental value of animals—they did not exist just to be used by Germans or anyone else (Arluke & Sax, 1992). They also frequently cited various emotions related to empathy in justifying their

⁶⁷ Evidence for empathy deficiency as a cause of bad behavior in general is also mixed. On violence, sexual abuse, and bullying see Maibom (2012) for a review. See Dadds (2008) for further discussion of predictors of animal abuse in childhood. See McPhedran (2009) for a critique of the relationship between empathy and cruelty similar in spirit to my own.

actions. For instance, one of their stated aims in promoting animal protection laws was to “awaken and strengthen compassion as one of the highest moral values of the German people” (Arluke & Sax, 1992, p. 8). Nazis are typically considered the archetype for cruelty and insensitivity, so it would seem that their cruel traits must either be caused by something unrelated to empathy or by a very specific form of empathy failure.

Another potential counterexample comes from work derived from Stanley Milgram’s famous shock experiments. In these experiments, participants were instructed to deliver increasingly painful shocks to someone who was ostensibly just another participant. Across a wide variety of different experimental conditions, people readily delivered the shocks, even when the target (who was actually a confederate enlisted by Milgram) explicitly asked not to continue and in some cases showed signs of harm, even apparent death. Sheridan and King (1972) conducted a variation on this experiment in which they replaced the human confederate with a puppy. However, for this experiment, the puppy received *actual* shocks. Sheridan and King found that 77% of participants were willing to deliver shocks to the puppy. Though the shocks were deemed to be harmless, they were jolting enough to occasionally elicit howls from the puppy. Apparently, this was not enough to motivate people to desist from following the instructions to shock.

One might affirm in response to these disturbing results that these participants were indeed failing to empathize with the puppy. A common explanation for Milgram’s results is that people readily conform to authority. In the case of the puppy, it could be said, people similarly conformed to authority and thus failed to consider the situation

from the puppy's perspective. The circumstances were so abnormal, and the external pressure so strong, that participants' empathy either wasn't engaged properly or was easily overwhelmed. Either way, it still indicates an empathy failure. As Miller (2009) argues, one can make a case for the importance of empathy by simply saying that *if* the right triggers are available for eliciting empathy, then people will act kindly and provide assistance. In the Milgram experiments the right triggers were not available, so people could not respond empathically, leading to cruel behaviors.

This interpretation seems hopeless, however, if empathy is supposed to be central to moral concern. Empathy with animals must be able to operate outside of situations in which it is relatively easy to empathize. Many industries in which animals are said to be treated cruelly, including in laboratories and on farms, possess precisely the sort of authoritative structure used in Milgram's experiments. If the failure to treat animals well in intensive animal agriculture, for instance, is a result of an empathy failure, then it would seem that empathy is too weak to rely on for many of our moral aims. Presumably cruelty caused by empathy failures would be pervasive throughout society.

5.2.5 Empathy Expands our Moral Concern to Outgroups

One task commonly attributed to empathy is expanding concern to those for whom we normally have difficulty showing concern. Members of outgroups, typically those from countries or ethnicities that are not our own, receive less attention and moral concern than people in our ingroup. Empathy is responsible for expanding our "moral circle" in a way that our normal treatment of outgroups cannot accomplish. For example,

research has found that empathy can induce positive attitudes towards normally stigmatized people—those who are considered low in warmth and competence (Batson, Chang, Orr, & Rowland, 2002; Batson et al., 1997). This includes people with AIDS, homeless people, drug addicts, and convicted murderers.

We might predict that empathy plays a similar function with respect to animals. Many animals live in our homes as members of our family, and are at the very least considered honorary members of our ingroup, but many others are clearly treated as outgroup members. Perhaps empathy can induce positive attitudes towards even outgroup animals, and potentially motivate increased moral concern. This possibility is nicely characterized by Robert Solomon, in describing our expanding moral sense:

[W]hat allows the circle to expand is not reason...but rather knowledge and understanding in the sense of coming to appreciate the situations and the circumstances in which other people and creatures find themselves. This requires what many theorists now call "empathy" or "feeling with"... We *learn* to empathize with others.... We *learn* to perceive chickens, cows, and warthogs as sentient beings with real emotions, and we learn to conceive of our uses of animals as a moral choice, not, first of all, because of any rational principles but because of our cultivated and expanded emotional awareness. (1999, p. 75-76).

Here Solomon emphasizes that acquiring empathy and moral sensitivity to others is a learning process. This is conceding, to a certain extent, that we are not naturally sensitive to the emotions of outgroup members. But this still allows for the use of empathy to extend our concern to outgroups (indeed it provides a reason to encourage explicit teaching of empathy).

While this is a noble goal, it faces significant obstacles in being realized in human psychology. We might be able to bypass issues concerning our natural empathic abilities by advocating explicit use of empathy (where this includes prompts and

reminders), but this does not mean that people will be capable of setting aside other psychological limitations they possess in empathizing with outgroups. Notwithstanding some of Batson's work, research on human-oriented empathy has generally found that empathy does *not* motivate help for outgroups. When people do help outgroups, they usually do so because of some sense of similarity to themselves (Stürmer, Snyder, Kropp, & Siem, 2006). When people are given the option to empathize with both ingroups and outgroups, they generally focus only on their ingroup (Brown, Bradley, and Lang, 2006). This is true even if other typical elicitors are manipulated. For example, Stürmer, Snyder, and Omoto (2005) found ingroup biases even if outgroup members were physically attractive and even if outgroup members had a serious illness in need of treatment. An explanation for this phenomenon offered by recent brain imaging experiments is that our empathic neural substrates are attuned *only* to ingroups (Gutsell & Inzlicht, 2010; Gutsell & Inzlicht, 2012; Xu, Zuo, Wang, & Han, 2009).

A related phenomenon in empathizing with animals, as discussed in section 3, is the effect of physical similarity to humans. A common finding in various surveys (Batt, 2009; Eddy, Gallup, & Povinelli, 1993) is that people attribute mental states to animals based roughly on phylogenetic proximity and physical similarity to humans. A similar finding has been found with empathy and moral concern. For instance, Plous (1993) showed participants pictures of a monkey, raccoon, pheasant, and bullfrog, and told them that each animal had been abused in certain ways. Skin conductance measurements showed increased activity in response to the animals' similarity to humans (see also Opatow, 1993). Westbury and Neumann (2008) similarly found that empathic emotional

responses to animals in abusive situations increased according to phylogenetic similarity (as measured by survey as well as skin conductance responses).

This bias for similarity to humans can also be seen in people's allocation of punishment for animal abuse. In one experiment (Allen et al., 2002), people read about abuse of a goose, monkey, possum, or goanna. They were then asked how much punishment they would give the transgressor. Those who scored highest in empathy gave out harsher punishments, and this was mediated by similarity to humans. So those high in empathy did indeed rush to protect some animals, but this was limited to species nearest to us, most notably the primates. Moreover, people were willing to allocate harsh punishments for abusing these animals. It's not clear from this experiment whether people were *too* harsh, but this is nonetheless worrisome in that it replicates the favoritism shown generally to human ingroups.

As Jesse Prinz (2011a) says about the limitations of empathy, "We can no more overcome its limits than we can ride a bicycle across the ocean; it is designed for local travel" (p. 229). The evidence cited here suggests that many animals are not considered part of any ingroup for humans. The only way in which empathy can assist animals is if they possess physical, functional, or phylogenetic similarity to humans, or if they are made honorary ingroup members through domestic companionship. This makes empathy useful in terms of prompting moral concern for some animals, but its scope is too limited to expand generalized moral concern to all animals.

5.2.6 Empathy Motivates Us to Act on our Moral Duties

A common idea is that empathy is helpful for motivating moral action. For instance, we might in fact view animals positively, and recognize that they are owed improved treatment, but fail to act appropriately because we are not sufficiently motivated. Consider factory farming: few would deny that factory farming is morally problematic, yet few take action to improve or abolish such practices. What is needed, one might say, is more empathy, to transform our moral judgments into appropriate action. Empathy assists us in meeting the moral ends we have set for ourselves. I reject this claim: empathy is a weak motivator, and when it does motivate, the results often run contrary to our moral aims.

In response to the arguments of the previous section, it could be maintained that empathy is a strong motivator of moral concern, so long as it is directed at ingroup members. For instance, someone could point out that dogs in the U.S. are generally considered ingroup members, yet dog abuse and abandonment is still quite high. Promoting empathy could increase moral concern expressed towards these animals and others like them. Farm animals may also be potential targets of ingroup empathy. Domestic companions and livestock, assuming they are sufficiently close to human ingroups, demand moral attention best served by empathy.

A good deal of the evidence in support of the motivational powers of empathy in humans comes from the work of Daniel Batson. One application of Batson's methods to animals comes from Berenguer (2007), who also found strong evidence for the motivational powers of empathy. Berenguer showed people one of two pictures, either a

row of trees that had been cut down or a bird covered in oil. People were then asked to either judge the pictures objectively or to imagine the feelings of the entity affected and how the event affected that entity's feelings. Participants' opinions were then solicited on how a global funding agency should spend its money. Results showed that those who were asked to imagine the feelings of the target (the empathy condition) wanted more money to be donated to environmental causes (for both targets). This provides pretty straightforward evidence that empathy can motivate prosocial behaviors toward animals. It's also important to note that the intention to help was generalized. The assistance participants were willing to provide went beyond the targets and was aimed at environmental causes in general. This indicates that empathy is capable of producing general prosocial behaviors (which is consistent with other research on the effects of empathy towards humans; Eisenberg, Fabes, & Spinrad, 2006; Eisenberg & Miller, 1987).

I discussed above how this sort of phenomenon could also be accounted for by moral anger (which could also explain the effects of participants' responses to something like a bird covered in oil). Here I'll pursue a different objection, directed at the strength of motivation provided by empathy. A common criticism of Batson's research is that empathy produces only superficial helping (Neuberg et al., 1997). While the forms of helping uncovered by Bereunger are important (such as donating money to help animals), the centrality thesis sets a high bar for what empathy should accomplish.

One illuminating study indicating that empathy is a relatively weak motivator comes from Krueger et al. (2013). They injected participants with oxytocin, a hormone

known to increase empathy, after which participants read about different transgressions involving one person causing serious physical harm to another person (e.g., a robber punching a gas station attendant). The results showed that oxytocin injections increased participants' perception of harm but did not have any effect on how much punishment they thought the transgressors deserved. Though this might be cited as evidence that empathy produces moral concern (perceiving greater harm), it also perfectly illustrates empathy's inability to motivate moral behavior, even when clearly influencing people's moral judgments. Participants could have acted on the greater harm they perceived by delivering greater punishment, but they did not.

By contrast, the experiments discussed earlier indicate that anger is highly motivating, especially in responding to perceived transgressions. Of the six empirical theses I have examined, the centrality of moral anger—not empathy—is perhaps clearest on the issue of motivation. When attempting to rectify injustices, responding to personal threats, or simply trying to achieve some highly desired moral goal, anger seems to be the best candidate for providing motivational power (Carver & Harmon-Jones, 2009).

This is important when considering that in much of the empathy literature (including Berenguer, 2007), participants are *prompted* to empathize. Participants could also be prompted with other, potentially more motivating, emotions, like anger. It's also possible that, because empathy functions to elicit other emotions, the process of empathy is motivational *only insofar* as it elicits other, more motivating, emotions (again, like anger; Batson et al., 2007). If other emotions are more motivational, either through direct

prompting or indirectly through empathy, then those should be the focus for ethicists looking to motivate moral behavior.

5.3 Ethical Implications of Empathy

In this section, I have evaluated six prominent empirical claims made about empathy and found all of them to be, to different degrees, problematic. These claims are often taken to illustrate the centrality of empathy in producing moral concern for animals—they offer reasons to think empathy is psychologically essential to showing moral concern for animals. In rejecting these claims, I have also proposed that other, non-empathic, emotions are more capable of producing moral concern for animals. I have focused in particular on moral anger. Moral anger is highly motivating and can be used to respond to transgressions against animals.

However, many animal ethicists may feel as if something has gone wrong. Empathy is the quintessential moral emotion, so my argument must be mistaken. Perhaps I have identified the wrong empirical claims, or perhaps I have failed to identify the conception of empathy favored by specific ethical theories, or perhaps I need to account for the possibility that future evidence would contradict the studies I have reviewed here. Rather than address a wide range of potential criticisms, I will return again to cognitive empathy, and our ability to learn to be more attuned to animals' emotions. This is where I suspect the most significant disagreements will arise. After addressing cognitive empathy, I will conclude with brief suggestions for animal ethicists in turning to non-empathic emotions.

5.3.1 Controlling Empathy

Consider the definition Lori Gruen (2009) provides for what she variously calls “entangled” or “engaged” empathy:

Engaged empathy is a process whereby individuals who are empathizing with the well-being of others first respond to the other’s condition (most likely, but not exclusively, by way of a pre-cognitive empathetic reaction), and then reflectively imagine themselves in the position of the other, and then make a judgment about how the conditions that the other finds herself in may contribute to her state of mind or impact upon her interests. These judgments will involve assessing the salient features of the situation and require that the empathizer seek to determine what information is pertinent to effectively empathize with the being in question. (pp. 29-30)

Gruen’s position accommodates emotional contagion but primarily requires a great deal of cognition. Reflective imagination, I suggested above, is used infrequently, and mostly as a failsafe. Consideration of interests, while certainly important for expressing moral concern for animals, is not part of empathy, as I understand it. But someone might object that engaged, or cognitive, empathy is what animal ethicists *should* focus on. Ethicists cannot rely on our automatic processes to be sufficient.

Two features of empathy that make this proposal particularly problematic are empathy’s ingroup bias and its weak motivational powers (the fifth and sixth hypotheses examined above). These are inherent in empathy and extremely difficult to modify. If we want to produce moral concern for outgroups—like animals—we need a different emotion.

Suppose for the sake of argument, however, that my rejection of the fifth and sixth hypotheses was mistaken. I will highlight a third feature that makes cognitive empathy ineffective and unhelpful: what is known as Just World Bias or Just World

Theory. Lerner and colleagues' seminal research (Lerner, 1980; Lerner & Miller, 1978; Lerner & Simmons, 1966) found that victims who people were otherwise inclined to help were blamed for their situation if their suffering could not be alleviated. The explanation offered by psychologists is that this phenomenon is the result of a bias towards thinking that the world is fundamentally just, commonly expressed in the psychological literature as "people get what they deserve and deserve what they get." For instance, when people are asked to imagine how it would feel to be homeless—a group of people who experience persistent suffering—people tend to *reduce* their expressed concern for the homeless (Kozak, Marsh, & Wegner, 2006). The explanation for this is that people assume that such persistent suffering would not occur in a just world. Since the world is just, they reason, anyone who lives in poor conditions, such as the homeless, deserve their plight.

This is relevant to animals because empathy requires people to look at animals' external conditions. Gruen's definition includes this feature, as does Elisa Aaltola's (2012, chapter 6), among animal ethicists. I do not know of any research directly testing the relationship between just world beliefs and animals, but we can predict that people's responses will vary according to the plight of the animal in question. For instance, a pig living in wretched conditions might be considered *appropriate* for the pig. Imagining

pigs' external conditions is likely to make people content with their plight and think they deserve nothing better.⁶⁸

In short, the just world bias predicts that people will *search* for a reason to blame someone for their condition, rather than respond to injustice or real harms. Positive empathic responses will thus be limited by the extent to which our just world beliefs lead us to blame animals for their situation. Importantly, it is unlikely that this can be overcome with cognitive control. The Just World Bias is a general psychological process predicted to be employed even when we attempt to control our initial responses. Thus, we are likely to judge animals to be responsible for their poor conditions, even when we step back to fully imagine what it would be like to be in their situation.⁶⁹

5.3.2 Concluding Thoughts

Given my analysis of the centrality thesis, we must conclude that empathy is not central to producing moral concern for animals. And given what I have just said about just world beliefs, we must also be skeptical of the ability of cognitive processes to make empathy more central. Cognitive perspective taking can indeed function as a failsafe under certain conditions, but this is not sufficient to justify giving empathy a privileged role in our moral theories.

⁶⁸ Asking someone to empathize with domestic pets, by contrast, is likely to produce greater concern, as pets in the West often live comfortable lives among people who already care for them.

⁶⁹ There is also some evidence that anger, or “moral outrage,” is absent in cases of “system justification” (Wakslak, Jost, Tyler, & Chen, 2007). This again illustrates the importance of moral anger, though there is not enough empirical evidence to compellingly argue that anger is capable of overcoming the Just World Bias.

Instead, ethicists should turn to the plethora of other moral emotions we possess towards animals. In doing this, one aspect of empathy that should be retained is the focus on animals' external conditions. Seeing how animals live evokes emotional responses and drives us to consider whether we are treating them rightly or wrongly. This includes not just moral anger, but all the other emotions we normally feel in response to moral transgressions—guilt, shame, disgust, and contempt, among others. Ethicists have been right to focus on our emotional responses to animal abuses, but it is these other emotions—not empathy—that have the potential to actually produce moral concern for animals.

6. CONCLUSION: TOWARDS PSYCHOLOGICAL PLAUSIBILITY

To this point my conclusions have been largely negative—identifying what animal ethicists should avoid. Here I will make more positive suggestions and sketch some strategies for constructing psychologically plausible theories in animal ethics. First, however, I will review the main claims I have made thus far. Then I will return to issues raised about psychological plausibility and propose strategies for pursuing social and political change.

6.1 Summary of Main Claims

Section 2 aimed to identify basic psychological constraints on ethical theories. I argued against Owen Flanagan's account of the psychological plausibility of ethical theories on the grounds that it fails to actually rule out any ethical theories, and then proposed five of my own criteria for assessing an ethical theory's psychological plausibility. These criteria demand that ethicists 1) look to psychological science to support their theories, 2) ensure that their theories' demands can ultimately be met by human beings, 3) promote ideas that can have an impact on currently existing human beings, 4) have a plan for modifying human psychology if current psychological profiles are inadequate, and 5) make moral psychological claims that are broadly consistent with their theories' main goals. I also discussed ways in which psychological constraints impact ethical prescriptions for theories that meet these basic criteria. My argument, in short, was that when individuals possess limited resources, such that moral obligations cannot be met, then individuals do not have the relevant moral obligations.

The aim of sections 3, 4, and 5 was to identify the most significant psychological limitations relevant to animal ethics. I focused on attributing mental states to animals, the role of animals as disgust elicitors, and the role of empathy in producing moral concern for animals. To evaluate the normative implications of these issues, I primarily drew from the Impact, Achievability, and Transition criteria outlined in section 2. That is, I examined whether theories in animal ethics could have an impact on current psychological profiles, given these psychological factors, or if they could not, whether there was a plausible transition plan for people to be able to meet these demands at some point in the future.

First I examined mentalizing, and argued that the empirical evidence supports the phenomenal account of mentalizing, which holds that animals are deemed morally considerable only if they are attributed phenomenal mental states. Agential states, by contrast, do not factor into moral considerability judgments. I also provided evidence that phenomenal mental states are attributed mainly to animals that look and act like human beings, thereby limiting the range of animals deemed morally considerable. This poses a challenge to ethicists, I argued, since the range of animals thought to be capable of experiencing phenomenal states is far beyond those that look and act like us. This additionally places limits on animals' moral significance, insofar as moral considerability is a prerequisite for moral significance. My ultimate conclusion in section 3 was that pure agential views—those that deny the relationship between phenomenal states and moral considerability—fail to satisfy the Impact, Achievability,

and Transition criteria, while theories that attribute sentience to a wide range of animals satisfy these criteria but still face significant psychological difficulties.

In section 4, I surveyed research indicating that animals cause disgust through two different mechanisms: core disgust and animal reminder disgust (or mortality salience). In short, animals activate systems designed to protect us from contaminants, which leads us to show aversion towards animals. The aversive reactions caused by animals, I argued, raise two further challenges: 1) for psychologists, explaining how we cope with animals, despite their disgust-evoking powers; 2) for ethicists, accounting for this aversion when arguing that animals deserve improved moral treatment. I dealt with the psychological problem by drawing from research on dehumanization, and proposed that we cope with animals, and our disgust responses to them, by attributing mental states that include positive evaluations but simultaneously “mark” animals as inferior beings. That is, animals are evaluated positively but in a way that indicates they are sub-human. In order to address the ethical issue, I turned again to the criteria of Impact, Achievability, and Transition. I granted to “disgust skeptics” that our disgust responses could be managed by changing laws and policies, thus satisfying my minimal conditions for psychological plausibility. However, I argued that these changes do not necessarily entail moral psychological change for individuals, and thus individuals who still lack the resources to overcome their disgust responses cannot be expected to meet various moral obligations to animals.

Section 5 focused on empathy towards animals. I examined six empirical claims commonly made in support of the idea that empathy is central to showing moral concern

for animals—what I called the centrality thesis. I found each of these claims to be problematic, and thus rejected the centrality thesis. I also reviewed evidence indicating that other moral emotions, particularly anger, are more central to showing moral concern for animals. Since I acknowledged that empathy *could* produce moral concern for animals, I did not measure empathy-based theories using my minimal criteria for psychological plausibility. Instead, I provided reasons to think that attempts to improve or enhance empathy in order to thereby produce moral concern would ultimately fail. My suggestion for ethicists looking to improve moral concern for animals was to turn to other moral emotional responses to animals.

Overall, I found that most theories in animal ethics either fail to meet my minimal criteria for psychological plausibility or face significant psychological obstacles (despite meeting the minimal criteria). The latter category of theories—those that are minimally plausible but face significant challenges from human psychology—will be the focus of the rest of the section. The main question for theories that are minimally plausible is whether the ideals they promote can be realized through social and political change. In short, these theories need to pursue a transition plan, along the lines of my discussion in section 2.

6.2 Managing Psychological Limitations

One of the chief issues raised in section 2 was the implication of limited resources for individuals' moral obligations. I argued that when individual resources are in fact limited, such that moral obligations cannot be met, then individuals do not have

the relevant moral obligations. The solution usually offered for limitations in individuals' resources, which I sketched briefly in section 2 and described in more detail in section 4 on the topic of disgust, is to pursue social and political change so that individuals have greater resources to work with. That is, the best way to overcome psychological constraints is through incremental change at the level of laws and policies. In this section, I set aside the issue of individuals' moral obligations and focus instead on societies' obligations to change laws and policies.

One interesting theme of the psychological processes I have focused on is the role of animal bodies. Animal bodies determine (and limit) phenomenal mentalizing, elicit disgust in human beings, and identify animals as outgroup members, thus making animals unlikely targets for human empathy. My suggestions will thus focus on changing our perceptions of animal bodies. I will propose strategies for handling problems raised by animal bodies, specifically those influenced by our Just World Bias and disgust toward animals. These suggestions are largely speculative, since the relevant research has yet to be conducted. My aim is to make these suggestions amenable to further empirical study and practical applications.

Two preliminary issues must be addressed before proceeding. There are generally two types of proposals for managing psychological limitations. A society could protect against potential harmful impacts of psychological processes, or it could attempt to change psychological processes so they no longer have the potential to cause harm. I will be discussing the latter, and essentially taking for granted that it could be achieved through ordinary democratic processes. However, there has been much debate

over the extent to which liberal societies (like the U.S.) can engage in either of these projects, specifically with respect to animals (Donaldson & Kymlicka, 2011; Garner, 2005; Smith, 2012). The reason usually cited for why liberal societies cannot pursue these strategies is that liberal societies embrace the doctrine of “exclusion of ideals” (Raz, 1986, pp. 134-136). That is, they do not explicitly endorse any particular moral ideal in constructing laws and policies. This is problematic with respect to animals because not everyone values animals in the same way, and in many cases people’s values conflict. Though providing basic protections for animals is now widely considered permissible (Donaldson & Kymlicka, 2011), endorsing laws and policies that aim to change people’s moral psychological responses to animals is not.

However, I will not address the reasons usually given against the latter project in any detail, nor will I provide a thorough defense of its legitimacy. Changing moral psychologies through laws and policies is the primary aim of many ethicists and political philosophers, even if the strategies for doing so are not straightforwardly permissible in normal democratic processes. Insofar as change is inevitable, and laws and policies will nudge changes relevant to animals in one direction or another, we must assume that governments and institutions will have some sort of influence on these outcomes, and can be evaluated accordingly. I intend for my suggestions to be useful for changing laws and policies, but I leave it to other ethicists and political philosophers—those whose theories require moral psychological change—to determine how my suggestions might be incorporated into normal democratic processes.

The second preliminary issue that must be addressed is how the motivation and societal resources for moral psychological change are supposed to come about if the population at large does not yet support the change (because they do not have the right moral psychological profile). Though I think this is a difficult problem that ethicists would need to address in order to enact certain policies, my suggestions will be aimed at cases in which moral psychological change is being instituted by individuals acting independently of the population as a whole (e.g., these policies would not be put through a voting procedure). The cases I have in mind are akin to what Thaler and Sunstein (2009) call “nudges,” where people’s decision frameworks are modified in order to steer their choices in particular directions. As Sunstein (2013) discusses, there are in fact offices of the U.S. government (like the Office of Information and Regulatory Affairs) that possess the authority and the resources to initiate nudges, particularly through policy change. These nudges are open to public inspection, but they do not require public support in order to be enacted, thus avoiding the problem of obtaining public support prior to pursuing moral psychological change.

Based on the research discussed in previous sections, two highly problematic features of current psychological profiles are our tendency to defend injustices toward nonhumans (the just world bias) and to be disgusted by nonhumans (core disgust/mortality salience). These influence many other processes (including our bias towards things that look and act like us), thus fundamentally impacting our relationships with animals. Any attempts to change moral psychology will likely need some general plan for managing these processes.

6.2.1 Disgust and Animal Bodies

With respect to disgust and mortality salience, the main challenge is managing our perceptions of animal bodies. The role of bodies in eliciting disgust in particular is what prompted Martha Nussbaum to urge solutions focused on “the body itself, and our anxieties about it” (2003, p. 424). But as I discussed in section 4, *focusing* on the body will likely be unhelpful because it will enhance disgust towards animals. Similarly, I do not think increased focus on the body will assist in overcoming any psychological biases with respect to mentalizing or empathy. Instead, what is needed are proposals for changing how we perceive animal bodies.

One possible source for proposals on perceiving gross physical features is the literature on managing racism. For instance, Kelly, Faucher, and Machery (2010) review evidence for what they call “the dissemination hypothesis” as well as the well-known “contact hypothesis” (Gaertner, Rust, Dovidio, Bachman, & Anastasio, 1994). These ask, roughly, if either increased knowledge of other races or increased contact with other races will reduce harmful racist biases. We can ask the same questions of our relationships with animals. For instance, in a recent book review, Jessica Pierce (2013) proposes a combination of both strategies:

I challenge any skeptic to spend a few weeks immersed in the now extensive literature on animal cognition, emotion, and prosocial behavior—and then spend a couple of weeks in the company of animals—and still come away with a sense that they are less sensitive, less intelligent, less socially attuned than humans. Our skeptic will certainly come away thinking that animals can't do everything we do.... But she will also come away with a new appreciation of our own limited capacities and a new respect for other forms of life—including, I would think, a sense that our wanton cruelty and disregard for others' lives and feelings is just plain wrong.

Applying this to policy change, perhaps certain national organizations could be induced to provide the necessary resources for people to learn more about animals and to come into greater contact with them. Zoos and other scientific organizations engage in some of this already, but Pierce is clearly suggesting something more involved.

In Kelly, Faucher, and Machery's review, they argue that the dissemination hypothesis is unsupported by the evidence (knowing more about other races has no effect on racism) but that the contact hypothesis has some support. The evidence, they argue, suggests that coming into contact with other races does indeed reduce racism, albeit only slightly and only in specific circumstances. I will not provide a thorough review of the literature, but the evidence on the dissemination hypothesis for animals suggests the same negative conclusion (Hazel, Signal, & Taylor, 2011; Heleski & Zanella, 2006; Jamieson et al., 2012). I will thus focus on the contact hypothesis as it applies to animals.

What is needed is a better understanding of the types of contact with animals that can modify our aversive responses. Unfortunately research on contact with animals has not addressed this issue in much detail. Typically researchers study the effects of introducing animals that are already viewed positively, like dogs and rabbits (Daly & Suggs, 2010). "Disgusting" animals in these studies are usually somewhat familiar to the participants, like fish and frogs. Moreover, researchers usually investigate what people learn about animals from handling them, rather than how aversive responses might be managed. Thus, we do not have a very good idea of how to avoid aversive and avoidant reactions to animals.

One exception to these trends is a study by Randler, Hummel, and Prokop (2012). They were able to reduce 11-13 year old children's disgust towards a wood louse, a snail, and a mouse by exposing the children to each animal, combined with basic instruction about proper handling in order to avoid causing harm. Another study of 11-12 year-old children found that disgust and fear towards toads, frogs, and salamander were greatly influenced by direct experience with each type of animal, indicating that more experience led to more favorable attitudes (Tomažič, 2011). Though only suggestive, these studies indicate what sorts of interventions, if pursued on a larger scale, might be able to reduce people's aversive responses to animals.

To optimize the chances of success, future interventions would need to utilize a wide variety of animals (e.g., mammals that are likely to elicit disgust), which might prove difficult to implement on a larger scale, and thus require a larger amount of resources. Researchers and policy makers would also have to take into account the possibility that the way people respond to handling animals in controlled settings (like a classroom) does not generalize to other situations. For instance, perhaps aversion is reduced by bringing disgusting animals into clean, controlled environments but returns to baseline when the animals are encountered in their normal environment. Future research should look into whether the contact hypothesis is more effective for some types of animal exposures than others.

Increased contact with animals might also have trickle-down effects. For instance, Morris, Knight, and Lesley (2012) found that familiarity with animals, including rodents, increased attributions of mental states, particularly emotions. Given

my discussion of the phenomenal account in section 3, we can expect that increased contact with an animal increases the probability of that animal being deemed morally considerable. Depending on the strength of this relationship, increased contact might also help in overcoming our bias towards animals that look and act like human beings. For instance, perhaps people are naturally inclined to deny phenomenal mental states to rats, thereby also denying them moral considerability, but this can be at least partially reversed with certain types of exposure and handling.

6.2.2 Justifying Animals' Living Conditions

While contact with animals has some promise for reducing core disgust responses, it is arguably less suited for reducing mortality salience or just world responses. Just world responses are typically elicited by living conditions human beings perceive as unjust, while mortality salience is caused by symbolic threats to human existence. It is unlikely that increased exposure to animals would alleviate either of these responses. Instead, I will focus on changing animals' living conditions.

The main problem raised by the just world response is that we are likely to justify poor living conditions as being appropriate for animals, and on those grounds deny them improved treatment. The simplest way of avoiding this is to change animals' living conditions. Just World Theory predicts that the just world response is elicited when animals' conditions are indicative of misery and suffering. Misery and suffering cannot exist in a just world, so their existence must be explained some other way—as appropriate for animals. Rather than attempt to modify the just world response, so that

people see misery and suffering as bad in themselves, I suggest that a more effective solution would be to circumvent the just world response entirely. This might be achieved by changing animals' conditions so they are not indicative of misery and suffering. For instance, small, dirty enclosures isolated from other members of an animal's species will likely elicit the just world response in ordinary human agents; these are the sorts of conditions human beings would find uncomfortable and seek to justify. This could be avoided, however, by providing enclosures and living conditions that people would associate with improved treatment. This would not necessarily lead people to think that animals deserve a higher standard of treatment, but it would prevent the elicitation of just world responses.

This might also have the side effect of diminishing animals' mortality salience. Haslam's research, discussed in section 4, seems to indicate that we feel more comfortable with animals' presence when we can anthropomorphize them in certain ways. Exaggerating their perceptual adeptness, for example, clearly demarcates them as separate from us. Animals' living conditions can influence this exaggerated anthropomorphism. As I understand the process of dehumanization as it relates to mortality salience, human beings are likely to be threatened by either dirty living conditions or excessively lavish living conditions. Both processes reduce the gap between humans and nonhumans: the former by bringing humans down and reminding them of their animal natures, the latter by bringing nonhumans up and threatening human uniqueness. What are needed are living conditions that avoid both of these possible outcomes. We should work to create living conditions that are indicative neither

of misery and suffering nor similarity to human beings. This might involve somewhat superficial changes in living conditions—more directly informed by human biases than animal welfare—but animals would, as a result, be evaluated without the taint of mortality salience.

6.3 Conclusion: Towards Psychological Plausibility in Animal Ethics

As I explained in section 2, ethicists and political philosophers have been untroubled by psychological limitations because they think any pertinent limitations can be overcome through incremental changes. Societies, like individuals, possess resource constraints, but those can be overcome incrementally (e.g., through decades of economic growth). In this section, I have described some general strategies for managing human psychology in order to meet moral goals and pursuing incremental change in animal ethics. The main psychological obstacles I identified are people's perception of animal bodies and living conditions. Future proposals will likely need to develop roughly along the lines I have outlined, though much more research is needed in order to develop more specific strategies for moral psychological change.

A key factor in improving the psychological plausibility of ethical theories is providing individuals with greater resources. Bringing people into certain types of contact with animals, for instance, gives them greater resources for overcoming aversion they might have towards animals. Though this may seem like a simple idea, I hope to have shown throughout the various ways in which ethical theories fail to recognize that significant psychological constraints on individuals' reactions to animals exist that

cannot be effectively addressed, even incrementally, without significant commitments of societal resources. Solving this problem, and providing individuals with the resources to meet their moral obligations (as described by the main theories in animal ethics), requires greater attention to empirical research into moral psychology.

REFERENCES

- Aaltola, E. (2012). *Animal suffering: Philosophy and culture*. New York: Palgrave MacMillan.
- Allen, M. W., Hunstone M, Waerstad J, Foy, E., Hobbins, T., Wikner, B., & Wirrel, J. (2002). Human-to-animal similarity and participant mood influence punishment recommendations for animal abusers. *Society and Animals*, 10, 267–284.
- Angantyr, M., Eklund, J., & Hansen, E. M. (2011). A comparison of empathy for humans and empathy for animals. *Anthrozoös*, 24, 369-377.
- Appiah, K. A. (2008). *Experiments in ethics*. Cambridge, MA: Harvard University Press.
- Arico, A., Fiala, B., Goldberg, R.F., & Nichols, S. (2011). The folk psychology of consciousness. *Mind & Language*, 26, 327–352.
- Arluke, A., & Sanders, C. R. (1996). *Regarding animals*. Philadelphia, PA: Temple University Press.
- Arluke, A. and Sax, B. (1992). Understanding Nazi animal protection and the holocaust. *Anthrozoös*, 5, 6–31.
- Arndt, J., Cook, A., Goldenberg, J. L., & Cox, C. R. (2007). Cancer and the threat of death: The cognitive dynamics of death thought suppression and its impact on behavioral health intentions. *Journal of Personality and Social Psychology*, 92, 12–29.
- Arrindell, W. A., Mulkens, S., Kok, J., & Vollenbroek, J. (1999). Disgust sensitivity and the sex difference in fears to common indigenous animals. *Behaviour Research and Therapy*, 37, 273-280.

- Ascione, F. (1992). Enhancing children's attitudes about the humane treatment of animals: Generalization to human-directed empathy. *Anthrozoös*, 5, 176–191.
- Ascione, F. (1999). The abuse of animals and human interpersonal violence: Making the connection. In F. R. Ascione & P. Arkow (Eds.), *Child abuse, domestic violence, and animal abuse: Linking the circles of compassion for preventions and intervention* (pp. 50-61). West Lafayette, IN: Purdue University Press.
- Ascione, F. (Ed.) (2008). *The international handbook of animal abuse and cruelty: Theory, research, and application*. Purdue, IN: Purdue University Press.
- Baron-Cohen, S. (2011). *The science of evil: On empathy and the origins of cruelty*. New York: Basic Books.
- Bastian, B., Laham, S. M., Wilson, S., Haslam, N., & Koval, P. (2011). Blaming, praising, and protecting our humanity: the implications of everyday dehumanization for judgments of moral status. *British Journal of Social Psychology*, 50, 469-483.
- Batson, C. D. (2009). These things called empathy: Eight related but distinct phenomena. In Jean Decety (Ed.), *The social neuroscience of empathy* (pp. 4-8). Cambridge, MA: MIT Press.
- Batson, C. D., Chang, J., Orr, R., & Rowland, J. (2002). Empathy, attitudes, and action: Can feeling for a member of a stigmatized group motivate one to help the group? *Personality and Social Psychology Bulletin*, 28, 1656–1666.

- Batson, C. D., Kennedy, C. L., Nord, L. A., Stocks, E. L., Fleming, D. A., & Marzette, C. M. (2007). Anger at unfairness: Is it moral outrage? *European Journal of Social Psychology*, 37, 1272–1285.
- Batson, C. D., Klein, T. R., Highberger, L., & Shaw, L. L. (1995). Immorality from empathy-induced altruism: When compassion and justice conflict. *Journal of Personality and Social Psychology*, 68, 1042-1054.
- Batson, C. D., Lishner, D. A., Cook, J., & Sawyer, S. (2005). Similarity and nurturance: Two possible sources of empathy for strangers. *Basic & Applied Social Psychology*, 27, 15-25.
- Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., Klein, L. L., & Highberger, L. (1997). Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of Personality and Social Psychology*, 72, 105–118.
- Batt, S. (2009). Human attitudes towards animals in relation to species similarity to humans: a multivariate approach. *Bioscience Horizons*, 2, 180-190.
- Beatson, R. M., & Halloran, M. J. (2007). Humans rule! The effects of creatureliness reminders, mortality salience and self-esteem on attitudes towards animals. *British Journal of Social Psychology*, 46, 619-632.
- Beatson, R. M., Loughnan, S., & Halloran, M. J. (2009). Attitudes toward animals: The effect of priming thoughts of human-animal similarities and mortality salience on the evaluation of companion animals. *Society and Animals*, 17, 72-89.
- Becker, E. (1971). *The birth and death of meaning* (2nd ed.). New York: Free Press.

- Becker, E. (1973). *The denial of death*. New York: Free Press.
- Becker, E. (1975). *Escape from evil*. New York: Free Press.
- Berenguer, J. (2007). The effect of empathy in proenvironmental attitudes and behaviors. *Environment and Behavior*, 39, 269-283.
- Berker, S. (2009). The normative insignificance of neuroscience. *Philosophy & Public Affairs*, 37, 293-329.
- Bermúdez, J. L. (2009). *Decision theory and rationality*. Oxford: Oxford University Press.
- Bernstein, M. (1998). *On moral considerability*. New York: Oxford University Press.
- Biro, S., & Leslie, A.M. (2007). Infants' perception of goal-directed actions: Development through cue-based bootstrapping. *Developmental Science*, 10, 379–398.
- Boccatto, G., Capozza, D., Falvo, R., Durante, F. (2008). The missing link: ingroup, outgroup, and the human species. *Social Cognition*, 26, 224–34.
- Broida, J. P., Tingley, L., Kimball, R., & Miele, J. (1993). Personality differences between pro and antivivisectionists. *Society & Animals*, 1, 129-144.
- Brown, L. M., Bradley, M. M., & Lang, P. J. (2006). Affective reactions to pictures of ingroup and outgroup members. *Biological Psychology*, 71, 303-311.
- Brownlee, K. (2010). Reasons and ideals. *Philosophical Studies*, 151, 433-444.
- Buckels, E., & Trapnell, P. (2013). Disgust facilitates outgroup dehumanization. *Group Processes Intergroup Relations*, 16, 771-780.
- Buckwalter, W. & Phelan, M. (2013). Function and feeling machines: A defense of the

- philosophical conception of subjective experience. *Philosophical Studies*, 166, 349-361.
- Burke, B. L., Martens, A., & Faucher, E. H. (2010). Two decades of terror management theory: A meta-analysis of mortality salience research. *Personality and Social Psychology Review*, 14, 155-195.
- Burris, C. T., & Rempel, J. K. (2004). "It's the end of the world as we know it": Threat and the spatial-symbolic self. *Journal of Personality and Social Psychology*, 86, 19-42.
- Carruthers, P. (1992). *The animals issue*. Cambridge: Cambridge University Press.
- Carver, C. S. (2001). Affect and the functional bases of behavior: On the dimensional structure of affective experience. *Journal of Social Psychology Review*, 5, 345-356.
- Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: Evidence and implications. *Psychological Bulletin*, 135, 183-204.
- Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, 67, 319-333.
- Casebeer, W. D. (2003). *Natural ethical facts*. Cambridge, MA: MIT Press.
- Cavalieri, P. and Singer, P. (1993). *The great ape project: Equality beyond humanity*. New York: St. Martin's Press.
- Chapman, H. A. & Anderson, A. K. (2013). Things rank and gross in nature: A review and synthesis of moral disgust. *Psychological Bulletin*, 139, 300-327.

- Chapman, H., Kim, D., Susskind, J., & Anderson, A. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science*, 323, 1222-1226.
- Clipsham, P. (forthcoming). Does empirical moral psychology rest on a mistake? *Philosophical Studies*.
- Cochrane, A. (2012). *Animal rights without liberation*. New York: Columbia University Press.
- Cohen, G. A. (2003). Facts and principles. *Philosophy and Public Affairs*, 31, 211–45.
- Coplan, A. (2011). Understanding empathy: Its features and effects. In A. Coplan and P. Goldie (Eds.), *Empathy: Philosophical and psychological perspectives* (pp. 3-18). Oxford: Oxford University Press.
- Coplan, A. & Goldie, P. (2011). *Empathy: Philosophical and psychological perspectives*. Oxford: Oxford University Press.
- Copp, D. (2008). ‘Ought’ implies ‘can’ and the derivation of the Principle of Alternate Possibilities. *Analysis*, 68, 67-75.
- Costello K, & Hodson G. (2010). Exploring the roots of dehumanization: the role of animal-human similarity in promoting immigrant humanization. *Group Processes Intergroup Relations*, 13, 3–22.
- Cox, C. R., Goldenberg, J. L., Pyszczynski, T., & Weise, D. (2007). Disgust, creatureliness and the accessibility of death-related thoughts. *European Journal of Social Psychology*. 37, 494-507.
- Curtis, V., Aunger, R. & Rabie, T. (2004). Evidence that disgust evolved to protect from risk of disease. *Proceedings of the Royal Society of London, Series B*, 271

(Supplement), 131e133.

Curtis, V., & Biran, A. (2001). Dirt, disgust, and disease. *Perspectives in Biology and*

Medicine, 44, 17-31.

Cushman, F., & Greene, J. (2012). Finding faults: How moral dilemmas illuminate

cognitive structure. *Social Neuroscience, 7*, 269-279.

Cushman, F., & Young, L. (2009). The psychology of dilemmas and the philosophy of

morality. *Ethical Theory and Moral Practice, 12*, 9–24.

D'Arms, J. & Jacobson, D. (2000). Sentiment and value. *Ethics, 110*, 722–48.

D'Arms, J. & Jacobson, D. (2003). The significance of recalcitrant emotions (or Anti-

QuasiJudgmentalism). In A. Hatzimoysis (Ed.), *Philosophy and the Emotions*

(pp. 127-145). Cambridge: Cambridge University Press.

D'Arms, J. & Jacobson, D. (2007). Sensibility theory and projectivism. In D. Copp

(Ed.), *The Oxford handbook of ethical theory* (pp. 186–218). New York: Oxford

University Press.

Dadds, M. R. (2008). Conduct problems and cruelty to animals in children: What is the

link? In F. R. Ascione (Ed.), *The international handbook of animal abuse and*

cruelty (pp. 111-132). West Lafayette, IN: Purdue University Press.

Daly, B., & Sugg, S. (2010). Teachers' experiences with humane education and animals

in the elementary classroom: implications for empathy development. *Journal of*

Moral Education, 39, 101-112.

- Darley, J. M., & Batson, C. D. (1973). From Jerusalem to Jericho: A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology*, 27, 100-108.
- Darwall, S. (1998). Empathy, sympathy, care. *Philosophical Studies*, 89, 261-282.
- Darwall, S. (2011). Being with. *The Southern Journal of Philosophy*, 49, 4-24.
- Das, E., Bushman, B. J., Bezemer, M. D., Kerkhof, P., & Vermeulen, I. E. (2009). How terrorism news reports increase prejudice against out-groups: A terror management account. *Journal of Experimental Social Psychology*, 45, 453–459.
- Davey, G. C. L. (1992). Some characteristics of individuals with fear of spiders. *Anxiety Research*, 4, 299-314.
- Davey, G. C. L. (1994). Self-reported fears to common indigenous animals in an adult UK population: the role of disgust sensitivity. *British Journal of Psychology*, 85, 541-554.
- Davey, G. C. L., McDonald, A. S., Hirisave, U., Prabhu, G.G., Iwawaki, S., Jim, C. I.,... Reimann, B. C. (1998). A cross-cultural study of animal fears. *Behaviour Research and Therapy*, 36, 735–750.
- Davey, G. C. L., & Marzillier, S. J. (2009). Disgust and animal phobias. In B. O. Olatunji and D. McKay (Eds.), *Disgust and its disorders: Theory, assessment, and treatment implications* (pp. 169-190). Washington, DC, US: American Psychological Association.
- Davidson, D. (1980). Intending. In *Essays on actions and events*. Oxford: Oxford University Press.

- Davidson, D. (1985). Rational animals. In E. LePore and B. McLaughlin (Eds.), *Actions and events: Perspectives on the philosophy of Donald Davidson* (pp. 473-480). Oxford: Blackwell.
- Davis, M. H., Mitchell, K. V., Hall, J. A., Lothert, J., Snapp, T., & Meyer, M. (1999). Empathy, expectations and situational preferences: Personality influences on the decision to participate in volunteer helping behaviors. *Journal of Personality*, 67, 469-503.
- de Vignemont, F., & T. Singer (2006). The empathic brain: How, when and why? *Trends in Cognitive Sciences*, 10, 435-441.
- Decety, J., Echols, S. & Correll, J. (2010). The blame game: The effect of responsibility and social stigma on empathy for pain. *Journal of Cognitive Neuroscience*, 22, 985-997.
- DeGrazia, D. (1996). *Taking animals seriously*. Cambridge, UK: Cambridge University Press.
- DeGrazia, D. (2008). Moral status as a matter of degree? *Southern Journal of Philosophy*, 46, 181-198.
- DesAutels, P. (1998). Psychologies of moral perceivers. *Midwest Studies in Philosophy*, 22, 266-279.
- Donaldson, S. & Kymlicka, W. (2011). *Zoopolis*. Oxford: Oxford University Press.
- Donovan, J. (1996). Attention to suffering: A feminist caring ethic for the treatment of animals. *Journal of Social Philosophy*, 27, 81-102.
- Doris, J. M. (1998). Persons, situations, and virtue ethics. *Noûs*, 32, 504-30.

- Doris, J. M. (2002). *Lack of character: personality and moral behavior*. New York: Cambridge University Press.
- Doris, J. M. (2010). Heated agreement: Lack of character as being for the good. *Philosophical Studies*, 148, 135-46.
- Doris, J. M. & Stich, S. P. (2005). As a matter of fact: Empirical perspectives on ethics. In F. Jackson and M. Smith (Eds.), *The Oxford handbook of contemporary philosophy* (pp. 114-152). Oxford: Oxford University Press.
- Dunkel, C. S. (2009). The association between thoughts of defecation and thoughts of death. *Death Studies*, 33, 356–371.
- Eddy, T.J., Gallup, G.G., Jr., & Povinelli, D.J. (1993). Attribution of cognitive states to animals: Anthropomorphism in comparative perspective. *Journal of Social Issues*, 49, 87-101.
- Eisenberg, N. (1988). Empathy and sympathy: A brief review of the concepts and empirical literature *Anthrozoös*, 2, 15-17.
- Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In W. Damon & R. Lerner (Eds.), *Handbook of child psychology*, Vol. 3, Social, Emotional, and Personality Development, 6th ed., (pp. 647–702). Hoboken, NJ: John Wiley & Sons.
- Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin*, 101, 91-119.

- Ellingsen, K., Zanella, A. J., Bjerkås, E., & Indrebø, A. (2010). The relationship between empathy, perception of pain and attitudes toward pets among Norwegian dog owners. *Anthrozoös*, 23, 231-243.
- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114, 864-886.
- Erlanger, A. C., & Tsytsarev, S. V. (2012). The relationship between empathy and personality in undergraduate students' attitudes toward nonhuman animal. *Society and Animals*, 20, 21-39.
- Eskine, K. J., Kacinik, N. A. & Prinz, J. J. (2011). A bad taste in the mouth: Gustatory disgust influences moral judgment. *Psychological Science*, 22, 295-299.
- Estlund, D. (2011). Human nature and the limits (if any) of political philosophy. *Philosophy and Public Affairs*, 39, 207-237.
- Farrelly, C. (2007). Justice in ideal theory: A refutation. *Political Studies*, 55, 844-864.
- Favre, D. S. (2011). *Animal law: Welfare interests and rights* (2nd ed.). New York: Aspen Publishers.
- Fessler, D. & Haley, K. (2006). Guarding the perimeter: The outside-inside dichotomy in disgust and bodily experience. *Cognition and Emotion*, 20, 3-19.
- Fessler, D. & Navarrete, C. (2003). Meat is good to taboo: Dietary prescriptions as a product of the interaction of psychological mechanisms and social processes. *Journal of Cognition and Culture*, 3, 1-40.
- Fessler, D. & Navarrete, C. (2005). The effect of age on death disgust: Challenges to terror management perspectives. *Evolutionary Psychology*, 3, 279-296.

- Filippi, M., Riccitelli, G., Falini, A., di Salle, F., Vuilleumier, P., Comi, G., & Rocca, M. A. (2010). The brain functional networks associated to human and animal suffering differ among omnivores, vegetarians and vegans. *PLoS ONE*, 5, 1-9.
- Flanagan, O. (1991). *Varieties of moral personality: Ethics and psychological realism*. Cambridge, MA: Harvard University Press.
- Fox, M. A., & McLean, L. (2008). Animals in moral space. In J. Castricano (Ed.) *Animal subjects: An ethics reader in a post-human world*. Waterloo, CA: Wilfried Laurier University Press.
- Francione, G. (2000). *Introduction to animal rights: Your child or your dog?* Philadelphia: Temple University Press.
- Frey, R. G. (1980). *Interests and rights: The case against animals*. Oxford: Clarendon Press.
- Frey R. G. (1987). Autonomy and the value of animal life. *The Monist*, 70, 49-63.
- Frey, R. G. (2014). Moral standing, the value of lives, and speciesism. In H. LaFollette (Ed.), *Ethics in practice* (4th ed.) (pp. 321-338). Malden, MA: Wiley.
- Furnham, A., McManus, C., & Scott, D. (2003). Personality, empathy and attitudes to animal welfare. *Anthrozoös*, 16, 135-146.
- Gaertner, S., Rust, M., Bachman, B., Dovidio, J., & Anastasio, P. (1994). The Contact Hypothesis: The role of a common ingroup identity on reducing intergroup bias. *Small Group Research*, 25, 224-249.

- Gailliot, M. T., Stillman, T. F., Schmeichel, B. J., Maner, J. K., & Plant, E. A. (2008). Mortality salience increases adherence to salient norms and values. *Personality and Social Psychology Bulletin*, 24, 993.
- Gallo, I. S., Keil, A., McCulloch, K. C., Rockstroh, B., & Gollwitzer, P. M. (2009). Strategic automation of emotion regulation. *Journal of Personality and Social Psychology*, 96, 11-31.
- Garner, R. (2005). *The political theory of animal rights*. Manchester, UK: Manchester University Press.
- Garner, R. (2013). *A theory of justice for animals: Animal rights in a nonideal world*. Oxford: Oxford University Press.
- Gheaus, A. (2013). The feasibility constraint on the concept of justice. *The Philosophical Quarterly*, 63, 443-464.
- Gilabert, P. (2011). Feasibility and socialism. *Journal of Political Philosophy*, 19, 52-63.
- Gilabert, P. & Lawford-Smith, H. (2012). Political feasibility: A conceptual exploration. *Political Studies*, 60, 809-825.
- Gil-White, F. J. (2001). Are ethnic groups biological 'species' to the human brain?: Essentialism in our cognition of some social categories. *Current Anthropology*, 42, 515-554.
- Goldenberg, J. L., Arndt, J., Hart, J., & Routledge, C. (2008). Uncovering an existential barrier to breast self-exam behavior. *Journal of Experimental Social Psychology*, 44, 260-274.

- Goldenberg, J. L., Pyszczynski, T., Greenberg, J., & Solomon, S. (2000). Fleeing the body: A terror management perspective on the problem of human corporeality. *Personality and Social Psychology Review*, 4, 200–218.
- Goldenberg, J. L., Cox, C. R., Pyszczynski, T., Greenberg, J., & Solomon, S. (2002). Understanding human ambivalence about sex: The effects of stripping sex of meaning. *Journal of Sex Research*, 39, 310–320.
- Goldenberg, J. L., Pyszczynski, T., McCoy, S. K., Greenberg, J., & Solomon, S. (1999). Death, sex, love, and neuroticism: Why is sex such a problem? *Journal of Personality and Social Psychology*, 77, 1173–1187.
- Goldenberg, J. L., Pyszczynski, T., Greenberg, J., Sheldon, S., Kluck, B., & Cornwell, R. (2001). I am *not* an animal: Mortality salience, disgust, and the denial of human creatureliness. *Journal of Experimental Psychology: General*, 130, 427–435.
- Goldman, A. (1993). Ethics and cognitive science. *Ethics* 103, 337–360.
- Goldman, A. (2011). Two routes to empathy: Insights from neuroscience. In A. Coplan & P. Goldie (Eds.), *Empathy: Philosophical and psychological perspectives*. Oxford: Oxford University Press.
- Goodin, R. E. (1995). Political ideals and political practice. *British Journal of Political Science* 25, 37–56.
- Goodpaster, K. E. (1978). On being morally considerable. *The Journal of Philosophy*, 75, 308–325.
- Graham, P. A. (2011). Ought and ability. *Philosophical Review*, 120, 337–382.

- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*, 1029-1046.
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101*, 366-385.
- Gray, H., Gray, K., Wegner, D.M. (2007). Dimensions of mind perception. *Science, 315*, 619.
- Greenberg, J., Pyszczynski, T. & Solomon, S. (1986). The causes and consequences of a need for self-esteem: a terror management theory. In R. F. Baumeister (Ed.), *Public self and private self* (pp. 189–212). New York, NY: Springer.
- Greenberg, J., Pyszczynski, T., Solomon, S., Simon, L., & Breus, M. (1994). Role of consciousness and accessibility of death-related thoughts in mortality salience effects. *Journal of Personality and Social Psychology, 67*, 627–637.
- Greenberg, J., Solomon, S. & Pyszczynski, T. (1997). Terror management theory of self-esteem and cultural worldviews: Empirical assessments and conceptual refinements. *Advances in Experimental Social Psychology, 29*, 61-139.
- Greenberg, J., Solomon, S., & Arndt, J. (2008). A basic but uniquely human motivation: Terror management. In J. Y. Shah and W. L. Gardner (Eds.), *Handbook of motivation science* (pp. 114-134). New York: Guilford.
- Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., and Cohen, J.D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*, 2105–2108.

- Greene, J. (2008). The secret joke of Kant's soul. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Volume 3: The neuroscience of morality* (pp. 35-79). Cambridge: MIT Press.
- Greene, J., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, *111*, 364–371.
- Gruen, L. (2004). Empathy and vegetarian commitments. In S. S. Sapontzi (Ed.), *Food for thought: The debate over eating meat* (pp. 284-293). New York: Prometheus.
- Gruen, L. (2009). Attending to nature: Empathetic engagement with the more than human world. *Ethics & the Environment*, *14*, 23-38.
- Gruen, L. (2012). Navigating difference (again): Animal ethics and entangled empathy. In G. R. Smulewics-Zucker (Ed.), *Strangers to nature: Animal lives and human ethics* (pp. 213-234). Lanham, MD: Lexington Books.
- Gutsell, J., & Inzlicht, M. (2010). Empathy constrained: Prejudice predicts reduced mental simulation of actions during observation of outgroups. *Journal of Experimental Social Psychology*, *46*, 841–45.
- Gutsell, J. N., & Inzlicht, M. (2012). Intergroup differences in the sharing of emotive states: Neural evidence of an empathy gap. *Social Cognitive Affective Neuroscience*, *7*, 596-603.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*, 814–834.

- Haidt, J. & Bjorklund, F. (2008). Social intuitionists answer six questions about moral psychology. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Volume 2: The cognitive science of morality: Intuition and diversity* (pp. 181-218). Cambridge: MIT Press.
- Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research, 20*, 98-116.
- Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus: Special Issue on Human Nature, 133*, 55-66.
- Haidt, J. & Kesebir, S. (2010). Morality. In S. Fiske, D. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology, 5th Edition* (pp. 797-832). Hoboken, NJ: Wiley.
- Haidt, J., Koller, S., & Dias, M. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology, 65*, 613-628.
- Haidt, J., McCauley, C. R., & Rozin, P. (1994). Individual differences in sensitivity to disgust: A scale sampling seven domains of disgust elicitors. *Personality and Individual Differences, 16*, 701-713.
- Hare, R. M. (1963). *Freedom and reason*. Oxford: Clarendon Press.
- Harman, G. (1999). Moral philosophy meets social psychology: Virtue ethics and the fundamental attribution error. *Proceedings of the Aristotelian Society, 99*, 315-331.
- Harman, G. (2009). Skepticism about character traits. *Journal of Ethics, 13*, 235-242.

- Harmon-Jones, E., Gable, P., & Peterson, C. (2010). The role of asymmetric frontal cortical activity in emotion-related phenomena: A review and update. *Biological Psychology*, 84, 451-462.
- Harmon-Jones, E., Sigelman, J.D., Bohlig, A., & Harmon-Jones, C. (2003). Anger, coping, and frontal cortical activity: the effect of coping potential on anger-induced left frontal activity. *Cognition and Emotion*, 17, 1–24.
- Harmon-Jones, E., Simon, L., Greenberg, J., Pyszczynski, T., Solomon, S., & McGregor, H. (1997). Terror management theory and self-esteem: Evidence that increased self-esteem reduces mortality salience effects. *Journal of Personality and Social Psychology*, 72, 24–36.
- Haslam N. (2006). Dehumanization: an integrative review. *Personality and Social Psychology Review*, 10, 252–264.
- Haslam, N., Kashima, Y., Loughnan, S., Shi, J., & Suitner, C. (2008). Subhuman, inhuman, and superhuman: Contrasting humans with nonhumans in three cultures. *Social Cognition*, 26, 248-258.
- Haslam, N. & Loughnan, S. (2014). Dehumanization and infrahumanization. *Annual Review of Psychology*, 65, 399-423.
- Hatfield, E., Rapson, R. L., & Le, Y. L. (2009). Emotional contagion and empathy. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 19-30). Cambridge, MA: MIT Press.

- Hayes, J., Schimel, J., Arndt, J., & Faucher, E. H. (2010). A theoretical and empirical review of the death-thought accessibility concept in terror management research. *Psychological Bulletin*, 136, 699–739.
- Hazel, S. J., Signal, T. D., & Taylor, N. (2011). Can teaching veterinary and animal-science students about animal welfare affect their attitude toward animals and human-related empathy? *Journal of Veterinary Medical Education*, 38, 74-83.
- Held, V. (2006). *The ethics of care: Personal, political, and global*. New York: Oxford University Press.
- Heleski, C. R., & Zanella, A. J. (2006). Animal science student attitudes to farm animal welfare. *Anthrozoös*, 19, 3-16.
- Henry, B. C. (2006). Empathy, home environment, and attitudes toward animals in relation to animal abuse. *Anthrozoös*, 19, 17-34.
- Heuer, U. (2010). Reasons and impossibility. *Philosophical Studies*, 147, 235-246.
- Herzog, H. (2010). *Some we love, some we hate, some we eat: Why it's so hard to think straight about animals*. New York: Harper Collins.
- Hills, A. M. (1993). The motivational bases of attitudes toward animals. *Society & Animals*, 1, 111-128.
- Hirschberger, G., Florian, V., & Mikulincer, M. (2005). Fear and compassion: A terror management analysis of emotional reactions to persons with disabilities. *Rehabilitation Psychology*, 50, 246–257.
- Holton, R. & Langton, R. (1999). Empathy and animal ethics. In D. Jamieson (Ed.), *Singer and his critics* (pp. 209-232). Oxford: Basil Blackwell.

- Horowitz, A. C. & Bekoff, M. (2007). Naturalizing anthropomorphism: Behavioral prompts to our humanizing of animals. *Anthrozoös*, 20, 23-35.
- Howe, D. (2013). *Empathy: What it is and why it matters*. New York: Palgrave Macmillan.
- Huebner, B. (2010a). Commonsense concepts of phenomenal consciousness: Does anyone care about functional zombies? *Phenomenology and Cognitive Science*, 9, 133-155.
- Huebner, B. (2010b, April). Commentary on T.J. Kasperbauer and David Wright's "Minding and liking nonhumans." Presentation at the Southern Society for Philosophy and Psychology (Atlanta, GA).
- Iacoboni, M. (2011). Within each other: Neural mechanisms for empathy in the primate brain. In A. Coplan & P. Goldie (Eds.), *Empathy: Philosophical and psychological perspectives* (pp. 45-57). Oxford: Oxford University Press.
- Ickes, W. (Ed.) (1997). *Empathic accuracy*. New York: Guilford Press.
- Jack, A. I. & Robbins, P. (2012). The phenomenal stance revisited. *Review of Philosophy and Psychology*, 3, 383-403.
- Jamieson, J., Reiss, M. J., Allen, D., Asher, L., Wathes, C. M., & Abeyesinghe, S. M. (2012). Measuring the success of a farm animal welfare education event. *Animal Welfare*, 21, 65-75.
- Jensen, M. (2009). The limits of practical possibility. *Journal of Political Philosophy*, 17, 168-84.

- Johnson, M. L. (1996). How moral psychology changes moral theory. In L. May, M. Friedman, & A. Clark (Eds.), *Mind and morals: Essays on cognitive science and ethics* (pp. 45–68). Cambridge, MA and London: MIT Press.
- Kamm, F. (2009). Neuroscience and moral reasoning: A note on recent research. *Philosophy & Public Affairs*, 21, 384-388.
- Kass, L. (1997). The wisdom of repugnance. *The New Republic*, 216, 17-26.
- Kelly, D. (2011). *The nature and moral significance of disgust*. Cambridge, MA: MIT Press.
- Kelly, D., Faucher, L., & Machery, E. (2010). Getting rid of racism: Assessing three proposals in light of psychological evidence. *Journal of Social Philosophy*, 41, 293-322.
- Kelly, D., & Morar, N. (forthcoming). Against the yuck factor. *Utilitas*.
- Kennett, J., & Fine, C. (2009). Will the real moral judgment please stand up? *Ethical Theory and Moral Practice*, 12, 77-96.
- Kirkpatrick, L. A. & Navarrete, C. D. (2006). Reports of my death anxiety have been greatly exaggerated: A critique of terror management theory from an evolutionary perspective. *Psychological Inquiry*, 17, 288-298.
- Klimecki, O., & Singer, T. (2013). Empathy from the perspective of social neuroscience. In A. Jorge and P. Vuilleumier (Eds.), *Cambridge handbook of human affective neuroscience* (pp. 533-549). Cambridge: Cambridge University Press.

- Knight, S., Bard, K., Vrij, A., & Brandon, D. (2010). Human rights, animal wrongs? Exploring attitudes toward animal use and possibilities for change. *Society and Animals, 18*, 251-272.
- Knobe, J. & Prinz, J. (2008). Intuitions about consciousness: Experimental studies. *Phenomenology and Cognitive Science, 7*, 67-83.
- Korsgaard, C. M. (1996). *The sources of normativity*. Cambridge: Cambridge University Press.
- Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology, 90*, 543-555.
- Krueger, F., Parasuraman, R., Moody, L., Twieg, P., de Visser, E., McCabe, K....Lee, M. R. (2013). Oxytocin selectively increases perceptions of harm for victims but not the desire to punish offenders of criminal offenses. *SCAN, 8*, 494-498.
- Kumar, V., & Campbell, R. (2012). On the normative significance of experimental moral psychology. *Philosophical Psychology, 25*, 311-330.
- Lamm, C., Batson, C. D. & Decety, J. (2007). The neural substrate of human empathy: Effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience, 19*, 42-58.
- Landau, M. J., Solomon, S., Greenberg, J., Cohen, F., Pyszczynski, T., Arndt, J., . . . Cook, A. (2004). Deliver us from evil: The effects of mortality salience and reminders of 9/11 on support for President George W. Bush. *Personality and Social Psychology Bulletin, 30*, 1136–1150.

- Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion*. New York: Plenum Press.
- Lerner, M. J. & Miller, D. T. (1978). Just world research and the attribution process: Looking back and ahead. *Psychological Bulletin*, 85, 1030-1051.
- Lerner, M. J., & Simmons, C. H. (1966). Observer's reaction to the "innocent victim": Compassion or rejection? *Journal of Personality and Social Psychology*, 4, 203–210.
- Levinson, B. M. (1978). Pets and personality development. *Psychological Reports*, 42, 1031-1038.
- Leyens J-P, Rodriguez-Torres, R., Rodriguez-Perez, A., Gaunt, R., Paladino, M., Vaes, J., & Demoulin, S. (2001). Psychological essentialism and the differential attribution of uniquely human emotions to ingroups and outgroups. *European Journal of Social Psychology*, 31, 395–411.
- Linzey, A. (Ed.). (2009). *The link between animal abuse and human violence*. Brighton and Portland: Sussex Academic Press.
- Luke, B. (1995). Taming ourselves or going feral? Toward a nonpatriarchal metaethic of animal liberation. In C. J. Adams and J. Donovan (Eds.), *Animals and women* (pp. 230-319). Durham: Duke University Press.
- Maibom, H. (2012). The many faces of empathy and their relation to prosocial action and aggression inhibition. *WIREs Cognitive Science*, 3, 253-263.
- Martens, A., Greenberg, J., Schimel, J., & Landau, M. J. (2004). Ageism and death: Effects of mortality salience and perceived similarity to elders on reactions to

- elderly people. *Personality and Social Psychology Bulletin*, 30, 1524–1536.
- Martins, Y. & Pliner, P. (2006). “Ugh! That’s disgusting!”: Identification of the characteristics of foods underlying rejections based on disgust. *Appetite*, 46, 75-85.
- Matchett, G., & Davey, G. C. L. (1991). A test of a disease-avoidance model of animal phobias. *Behaviour Research and Therapy*, 29, 91-94
- McGinn, C. (2011). *The meaning of disgust*. New York: Oxford University Press.
- McPhedran, S. (2009). A review of the evidence for associations between empathy, violence, and animal cruelty. *Aggression and Violent Behavior*, 14, 1-4.
- McShane, K. (2007). Why environmental ethics shouldn’t give up on intrinsic value. *Environmental Ethics*, 29, 43-61.
- McShane, K. (2011). Neosentimentalism and environmental ethics. *Environmental Ethics*, 33, 5-23.
- Mele, A. (2006). *Free will and luck*. New York: Oxford University Press.
- Merritt, M. (2000). Virtue ethics and situationist personality psychology. *Ethical Theory and Moral Practice*, 3, 365-383.
- Milgram, S. (1974). *Obedience to authority*. New York, NY: Harper and Row.
- Miller, C. (2009). Empathy, social psychology, and global helping traits. *Philosophical Studies*, 142, 247-275.
- Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *NeuroImage*, 28, 757-762.

- Morris, P., Knight, S., & Lesley, S. (2012). Belief in animal mind: Does familiarity with animals influence beliefs about animal emotions. *Society & Animals*, 20, 211-224.
- Munro, L. (2005). *Confronting cruelty*. Leiden and Boston: Brill.
- Murphy, L. (2000). *Moral demands in nonideal theory*. Oxford: Oxford University Press.
- Navarrete, C. D. (2005). Normative bias and adaptive challenges: A relational approach to coalitional psychology and a critique of terror management theory. *Evolutionary Psychology*, 3, 297-325.
- Navarrete, C. D. & Fessler, D. M. T. (2006). Disease avoidance and ethnocentrism: the effects of disease vulnerability and disgust sensitivity on intergroup attitudes. *Evolution and Human Behavior*, 27, 270-282.
- Neuberg, S. L., Cialdini, R. B., Brown, S. L., Luce, C., Sagarin, B. J., & Lewis, B. P. (1997). Does empathy lead to anything more than superficial helping? Comment on Batson et al. (1997). *Journal of Personality and Social Psychology*, 73, 510-16.
- Nussbaum, M. (2003). *Upheavals of thought: The intelligence of the emotions*. Cambridge, UK: Cambridge University Press.
- Nussbaum, M. (2004a). Beyond 'Compassion and Humanity': Justice for nonhuman animals. In C. R. Sunstein and M. C. Nussbaum (Eds.), *Animal rights: Current debates and new directions* (pp. 299–320). Oxford, UK: Oxford University Press.

- Nussbaum, M. (2004b). *Hiding from humanity: Disgust, shame, and the law*. Princeton, NJ: Princeton University Press.
- Nussbaum, M. (2006). *Frontiers of justice: Disability, nationality, species membership*. Cambridge, MA: The Belknap Press.
- Nussbaum, M. (2010). *From disgust to humanity: Sexual orientation and constitutional law*. Oxford, UK: Oxford University Press.
- Oaten, M., Stevensen, R. J., & Case, T. I. (2011). Disease avoidance as a functional basis for stigmatization. *Philosophical Transactions of the Royal Society B*, 366, 3433-3452.
- Olatunji, B. O., Moretz, M., McKay, D., Bjorklund, F., de Jong, P., Haidt, J. ... Schienle, A. (2009). Confirming the three-factor structure of the disgust scale-revised in eight countries. *Journal of Cross-Cultural Psychology*, 40, 234-255.
- Olatunji, B. O., Haidt, J., McKay, D., & David, B. (2008). Core, animal reminder, and contamination disgust: Three kinds of disgust with distinct personality, behavioral, physiological, and clinical correlates. *Journal of Research in Personality*, 42, 1243-1259.
- Opatow, S. V. (1993). Animals and the scope of justice. *Journal of Social Issues*, 49, 71-85.
- Oxley, J. C. (2011). *The moral dimensions of empathy*. New York: Palgrave Macmillan.
- Palmer, C. (2010). *Animal ethics in context*. New York: Columbia University Press.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.

- Patterson-Kane, E. G. & Piper, H. (2009). Animal abuse as a sentinel for human violence: A critique. *Journal of Social Issues*, 65, 589-614.
- Paul, E. S. (2000). Empathy with animals and with humans: Are they linked? *Anthrozoös*, 13, 194-202.
- Phillips, C. J. C., & S. McCulloch (2005). Student attitudes on animal sentience and use of animals in society. *Journal of Biological Education*, 40, 17-24.
- Pierce, J. (2013, December 28). [Review of the book *Can animals be moral?*, by M. Rowlands]. *Notre Dame Philosophical Reviews*. Retrieved from <http://ndpr.nd.edu/news/45240-can-animals-be-moral/>
- Pilley, J. W. & Reid, A. K. (2011). Border collie comprehends object names as verbal referents. *Behavioural Processes*, 86, 184-195.
- Plous, S. (1993). Psychological mechanisms in the human use of animals. *Journal of Social Issues*, 49, 11-52.
- Preston, S. D., & De Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25, 1-20.
- Preston, S. D., & Hofelich, A. J. (2012). The many faces of empathy: Parsing empathic phenomena through a proximate, dynamic-systems view of representing the other in the self. *Emotion Review*, 4, 24–33.
- Preylo, B. D., & Arikawa, H. (2008). Comparison of vegetarians and non-vegetarians on pet attitude and empathy. *Anthrozoös*, 21, 387-395.
- Prinz, J. (2007). *The emotional construction of morals*. Oxford: Oxford University Press.
- Prinz, J. (2011a). Against empathy. *Southern Journal of Philosophy*, 49, 214–233.

- Prinz, J. (2011b). Is empathy necessary for morality? In A. Coplan and P. Goldie (Eds.), *Empathy: Philosophical and psychological perspectives* (pp. 519–38). Oxford: Oxford University Press.
- Prokop, P., & Fančovičová, J. (2011). The effect of owning animals on perceived vulnerability to, and avoidance of, parasitic diseases in humans. *Journal of Individual Differences*, 32, 129–136.
- Prokop, P., Fančovičová, J., & Fedor, P. (2010). Health is associated with antiparasite behavior and fear of disease-relevant animals in humans. *Ecological Psychology*, 22, 222–237.
- Rachels, J. (2011). *The elements of moral philosophy* (7th ed.). New York: McGraw-Hill.
- Randler, C., Hummel, E., & Prokop, P. (2012). Practical work at school reduces disgust and fear of unpopular animals. *Society & Animals*, 20, 61-74.
- Rawls, J. (1999). *Law of peoples*. Cambridge, MA: Harvard University Press
- Rawls, J. (2001) *Justice as fairness: A restatement*. Cambridge, MA: Harvard University Press.
- Regan, T. (1983). *The case for animal rights*. Berkeley: University of California Press.
- Regan, T. (1991). *The thee generation*. Philadelphia, PA: Temple University Press.
- Robbins, P. (2008). Consciousness and the social mind. *Cognitive Systems Research*, 9, 15–23.
- Robbins, P. & Jack, A. (2006). The phenomenal stance. *Philosophical Studies*, 127, 59-85.
- Rosenblatt, A., Greenberg, J., Solomon, S., Pyszczynski, T., & Lyon, D. (1989).

- Evidence for terror management theory: I. The effects of mortality salience on reactions to those who violate or uphold cultural values. *Journal of Personality and Social Psychology*, 57, 681–690.
- Routledge, C., Ostafin, B., Juhl, J., Sedikides, C., Cathey, C., & Liao, J. (2010). Adjusting to death: The effects of mortality salience and self-esteem on psychological well-being, growth motivation, and maladaptive behavior. *Journal of Personality and Social Psychology*, 99, 897-916.
- Rozin, P. (2008). Hedonic “adaptation”: Specific habituation to disgust/death elicitors as a result of dissecting a cadaver. *Judgment and Decision Making*, 3, 191-194.
- Rozin, P. & Fallon, A. E. (1987). A perspective on disgust. *Psychological Review*, 94, 23-41.
- Rozin, P., Haidt, J., & Fincher, K. (2009). From oral to moral. *Science*, 323, 1179-1180.
- Rozin, P., Haidt, J., & McCauley, C. (2000). Disgust. In M. Lewis & J. Haviland-Jones (Eds.) *Handbook of emotions*, 2nd ed. New York: Guilford Press.
- Rozin, P., Haidt, J., & McCauley, C. R. (2008). Disgust. In M. Lewis, J. Haviland, & L. F. Barrett (Eds.). *Handbook of emotions*, 3rd ed. (pp. 757-776). New York: Guilford Press.
- Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, 76, 574-586.

- Rozin, P., Markwith, M., & Stoess, C. (1997). Moralization and becoming a vegetarian. *Psychological Science*, 8, 67-73.
- Rozin, P., Millman, L., & Nemeroff, C. (1986). Operation of the laws of sympathetic magic in disgust and other domains. *Journal of Personality and Social Psychology*, 50, 703–712.
- Rozin, P., Nemeroff, C., Horowitz, M., Gordon, B., & Voet, W. (1995). The borders of the self: Contamination sensitivity and potency of the body apertures and other body parts. *Journal of Research in Personality*, 29, 318-340.
- Russell, P. S., & Giner-Sorolla, R. (2011). Moral anger is more flexible than moral disgust. *Social Psychological and Personality Science*, 2, 360–364.
- Russell, P. S. & Giner-Sorolla, R. (2013). Bodily-moral disgust: What it is, how it is different from anger and why it is an unreasoned emotion. *Psychological Bulletin*, 139, 328-351.
- Sabini, J. & Silver, M. (2005). Lack of character? Situationism critiqued. *Ethics* 115, 535-562.
- Sandler, R. (2003). The external goods approach to environmental virtue ethics. *Environmental Ethics*, 25, 279.
- Sandler, R. (2004). Towards an adequate environmental virtue ethic. *Environmental Values*, 13, 477-495.
- Sandler, R. (2007). *Character and environment: A virtue-oriented approach to environmental ethics*. New York: Columbia University Press.

- Sauer, H. (2011). Social intuitionism and the psychology of moral reasoning. *Philosophy Compass*, 6, 708-721.
- Sauer, H. (2012a). Morally irrelevant factors: What's left of the dual process-model of moral cognition? *Philosophical Psychology*, 25, 783-811.
- Sauer, H. (2012b). Psychopaths and filthy desks: Are emotions necessary and sufficient for moral judgment? *Ethical Theory and Moral Practice*, 15, 95-115.
- Sax, B. (2000). *Animals in the Third Reich*. New York: Continuum.
- Savage-Rumbaugh, E.S., Shanker, S. G., & Taylor, T. J. (2001). *Apes, language, and the human mind*. Oxford: Oxford University Press.
- Schaich Borg, J., Lieberman, D., & Kiehl, K. A. (2008). Infection, incest, and iniquity: Investigating the neural correlates of disgust and morality. *Journal of Cognitive Neuroscience*, 20, 1529–1546.
- Schaller, M. (2006). Parasites, behavioral defenses, and the social psychological mechanisms through which cultures are evoked. *Psychological Inquiry*, 17, 96–101.
- Schaller, M., & Duncan, L. A. (2007). The behavioral immune system: Its evolution and social psychological implications. In J. P. Forgas, M. G. Haselton, & W. von Hippel (Eds.), *Evolution and the social mind* (pp. 293–307). New York, NY: Psychology Press.
- Schaller, M., & Murray, D. R. (2008). Pathogens, personality, and culture: Disease prevalence predicts worldwide variability in sociosexuality, extraversion, and

- openness to experience. *Journal of Personality and Social Psychology*, 95, 212–221.
- Schmeichel, B. Gaillot, M. T., Filardo, E., McGregor, I., Gitter, S., & Baumeister, R. (2009). Terror management theory & self-esteem revisited: The roles of implicit self-esteem in mortality salience effects. *Journal of Personality & Social Psychology*, 96, 1077-1087.
- Schmeichel, B. J. & Martens, A. (2005). Self-affirmation and mortality salience: Affirming values reduces worldview defense and death-thought accessibility. *Personality and Social Psychology Bulletin*, 31, 658.
- Schnall, S., Haidt, J., Clore, G. & Jordan, A. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin*, 34, 1096-1109.
- Schultz, P. W. (2000). Empathizing with nature: The effects of perspective taking on concern for environmental issues. *Journal of Social Issues*, 56, 391-406.
- Segerdahl, P., Fields, W. M., & Savage-Rumbaugh, S. (2005). *Kanzi's primal language: The cultural initiation of primates into language*. New York: Palgrave Macmillan.
- Sen, A. (2006). What do we want from a theory of justice? *Journal of Philosophy*, 103, 215–38.
- Sen, A. (2009). *The idea of justice*. London: Allen Lane/Penguin.
- Serpell, J. (1996). *In the company of animals* (2nd ed.). Cambridge, UK. Cambridge University Press.

- Sevillano, V., Aragonés, J. I., & Schultz, P. W. (2007). Perspective taking, environmental concern, and the moderating role of dispositional empathy. *Environment and Behavior*, 39, 685-705.
- Shapiro, K. (1994). The caring sleuth: Portrait of an animal rights activist. *Society and Animals*, 2, 145-165.
- Shelton, M. L., & Rogers, R. W. (1981). Fear-arousing and empathy-arousing appeals to help: The pathos of persuasion. *Journal of Applied Social Psychology*, 11, 366-378.
- Sheridan, C.L., & King, R. G. (1972). Obedience to authority with an authentic victim. *Proceedings of the Annual Convention of the American Psychological Association*, 80, 165-166.
- Shweder, R. A., Much, N. C., Mahapatra, M. & Park, L. (1997). The ‘big three’ of morality (autonomy, community, and divinity), and the ‘big three’ explanations of suffering. In A. Brandt & P. Rozin (Eds.), *Morality and health*. New York: Routledge.
- Signal, T., & Taylor, N. (2007). Attitude to animals and empathy: Comparing animal protection and general community samples. *Anthrozoös*, 20, 125-130.
- Simmons, A. J. (2010). Ideal and nonideal theory. *Philosophy and Public Affairs*, 38, 5-36.
- Singer, P. (1981). *The expanding circle: Ethics and sociobiology*. New York: Farrar, Straus and Giroux.
- Singer, P. (2012). *Practical ethics* (3rd ed.). Princeton: Princeton University Press.

- Singer, T., Seymour, B., O'Doherty, J., Stephan, K., Dolan, R., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303, 1157-1162.
- Sinnott-Armstrong, W. (1984). 'Ought' conversationally implies 'Can', *Philosophical Review*, 93, 249–261.
- Sinnott-Armstrong, W. (2006). Moral intuitionism meets empirical psychology. In M. Timmons & T. Horgan (Eds.), *Metaethics after Moore* (pp. 339–365). New York: Oxford University Press.
- Slote, M. (2007). *The ethics of care and empathy*. London: Routledge Press.
- Slote, M. (2010). *Moral sentimentalism*. Oxford: Oxford University Press.
- Smith, K. (2012). *Governing animals*. Oxford: Oxford University Press.
- Smits, J. A. J., Telch, M. J., & Randall, P. K. (2002). An examination of the decline in fear and disgust during exposure-based treatment. *Behaviour Research and Therapy*, 40, 1243–1253.
- Solomon, R. C. (1999). Peter Singer's *Expanding Circle*: Compassion and the liberation of ethics. In D. Jamieson (Ed.), *Singer and his critics* (pp. 64-84). Oxford and Malden, MA: Blackwell.
- Solomon, S., Greenberg, J. & Pyszczynski, T. (1991). A terror management theory of social behavior: the psychological functions of self-esteem and cultural worldviews. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, vol. 24 (pp. 91–159). San Diego, CA: Academic Press.

- Solomon, S., Greenberg, J., & Pyszczynski, T. (2004). The cultural animal: Twenty years of terror management theory and research. In J. Greenberg, S. L. Koole, & T. Pyszczynski (Eds.), *Handbook of experimental existential psychology* (pp. 13–34). New York: Guilford.
- Steiner, G. (2008). *Animals and the moral community: Mental life, moral status, and kinship*. New York: Columbia University Press.
- Steiner, G. (2013). *Animals and the limits of postmodernism*. New York: Columbia University Press.
- Stemplowska, Z. (2008). What's ideal about ideal theory? *Social Theory and Practice*, 34, 319-340.
- Stocker, M. (1971). 'Ought' and 'Can.' *Australasian Journal of Philosophy*, 49, 303–316.
- Stocker, M. (1976). The schizophrenia of modern ethical theories. *The Journal of Philosophy*, 73, 453-466.
- Streumer, B. (2007). Reasons and impossibility. *Philosophical Studies*, 136, 351–384.
- Stueber, K. (2006). *Rediscovering empathy: Agency, folk psychology, and the human sciences*. Cambridge, MA: MIT Press.
- Stueber, K. (2012). Varieties of empathy, neuroscience and the narrativist challenge to the contemporary theory of mind debate. *Emotion Review*, 4, 55-63.
- Stürmer, S., Snyder, M., Kropp, A., & Siem, B. (2006). Empathy-motivated helping: The moderating role of group membership. *Personality and Social Psychology Bulletin*, 32, 943-956.

- Stürmer, S., Snyder, M., & Omoto, A. M. (2005). Prosocial emotions and helping: The moderating role of group membership. *Journal of Personality and Social Psychology, 88*, 532-546.
- Sunstein, C. (2013). *Simpler: The future of government*. New York: Simon & Schuster.
- Sytsma, J. & Machery, E. (2010). Two conceptions of subjective experience. *Philosophical Studies, 151*, 299-327.
- Sytsma, J. & Machery, E. (2012). The two sources of moral standing. *Review of Philosophy and Psychology, 3*, 303-324.
- Swift, A. (2008). The value of philosophy in nonideal circumstances. *Social Theory and Practice, 34*, 363-387.
- Tagar, M. R., Federico, C. M. & Halperin, E. (2011). The positive effect of negative emotions in protracted conflict: The case of anger. *Journal of Experimental Social Psychology, 47*, 157-164.
- Tam, K-P. (2013). Dispositional empathy with nature. *Journal of Environmental Psychology, 35*, 92–104.
- Taylor, N. & Signal, T. D. (2005). Empathy and attitudes towards animals. *Anthrozoös, 18*, 18–27.
- Taylor, P. (1986). *Respect for nature*. Princeton, NJ: Princeton University Press.
- Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. New York: Penguin.
- Thompson, K. L. & Gullone, E. (2003). Promotion of empathy and prosocial behavior in children through humane education. *Australian Psychologist, 38*, 175–182

- Tom, E. M., Duncan, I. J. H., Widowski, T. M., Bateman, K. G. & Leslie, K. E. (2002). Effects of tail docking using a rubber ring with or without anesthetic on behavior and production of lactating cows. *Journal of Dairy Science*, 85, 2257-2265.
- Tomažič, I. (2011). Seventh graders' direct experience with, and feelings toward, amphibians and some other nonhuman animals. *Society & Animals*, 19, 225-247.
- Tversky, A. & Kahneman D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453-458.
- Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business*, 59, 251-278.
- Tybur, J. M., Griskevicius, V., & Lieberman, D. (2009). Microbes, mating, and morality: Individual differences in three functional domains of disgust. *Personality Processes and Individual Differences*, 97, 103-122.
- Tybur, J. M., Lieberman, D., Kurzban, R., & DeScioli, P. (2013). Disgust: Evolved function and structure. *Psychological Review*, 120, 65-84.
- Valentini, L. (2009). On the apparent paradox of ideal theory. *Journal of Political Philosophy*, 17, 332-355.
- Van Leeuwen, F. & Park, J. H. (2009). Perceptions of social dangers, moral foundations, and political orientation. *Personality and Individual Differences*, 47, 169–173.
- Varner, G. (1998). *In nature's interests? Interests, animal rights and environmental ethics*. New York: Oxford University Press.
- Varner, G. (2012). *Personhood, ethics, and animal cognition: Situating animals in the two-level utilitarianism of R.M. Hare*. Oxford: Oxford University Press.

- Velleman, D. J. (2002). Motivation by ideal. *Philosophical Explorations*, 5, 89-103.
- Velleman, D. J. (2013). *Foundations for moral relativism*. Cambridge, UK: Open Book Publishers.
- Vranas, P. (2007). I ought, therefore I can. *Philosophical Studies*, 136, 167 – 216.
- Wakslak, C. J., Jost, J. T., Tyler, T. R., & Chen, E. S. (2007). Moral outrage mediates the dampening effect of system justification on support for redistributive social policies. *Psychological Science*, 18, 267-274.
- Ware, J., Jain, K., Burgess, I., & Davey, G.C.L. (1994). Disease-avoidance model: Factor analysis of common animals fears. *Behaviour Research and Therapy*, 32, 57–63.
- Warren, M. A. (2000). *Moral status*. New York: Oxford University Press.
- Webb, K. & Davey, G. C. L. (1992). Disgust sensitivity and fear of animals: Effect of exposure to violent or revulsive material. *Anxiety, Stress & Coping*, 5, 329-335.
- Wemelsfelder, F. (1999). The problem of animal subjectivity. In F. Dolins (Ed.), *Attitudes to animals* (pp. 37-53). Cambridge: Cambridge University Press.
- Westbury, H. R. & Neumann, D. L. (2008). Empathy-related responses to moving film stimuli depicting human and non-human animal targets in negative circumstances. *Biological Psychology*, 78, 66-74.
- Xu, X., Zuo, X., Wang, X., & Han, S. (2009). Do you feel my pain? Racial group membership modulates empathic neural responses. *Journal of Neuroscience*, 29, 8525–8529.

Yamada, M. & Decety, J. (2009). Unconscious affective processing and empathy: an investigation of subliminal priming on the detection of painful facial expressions. *Pain, 143*, 71-75.